

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2005, 04:24:42 ; Search time 1057 Seconds

(without alignments)
16356.258 Million cell updates/sec

Title: US-10-647-057-8

Perfect score: 9726

Sequence: 1 atgcgcggcgcacaaataa.....gagaaaaaggagaaataatga 9726

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9726	100.0	9726	3	US-09-841-786-8
2	9726	100.0	11130	3	US-09-841-786-15
3	2780	28.6	2780	3	US-09-841-786-10
4	2322	23.2	2322	3	US-09-841-786-13
5	2141	22.0	2141	3	US-09-841-786-11
6	1887	19.4	1887	3	US-09-841-786-12
7	1130	11.6	1130	3	US-09-841-786-9
8	116.2	1.2	5361	3	US-08-973-462-2
9	116.2	1.2	6152	3	US-08-973-462-1
10	87.6	0.9	7218	2	US-08-232-463-14
11	79.2	0.8	8155	3	US-08-956-171B-63
12	79.2	0.8	8155	3	US-08-956-171B-63
13	76	0.8	1039	3	US-08-781-986A-63
14	64.8	0.7	1891	3	US-09-902-540-1280
15	64.6	0.7	3095	9	5231168-1
16	63.4	0.7	612	3	US-09-902-540-1357
17	62.4	0.6	50000	3	US-09-662-254B-24
18	62.2	0.6	4766	6	PCT-US93-07261-10
19	62	0.6	767677	3	US-09-949-016-12147
20	62	0.6	767677	3	US-09-949-016-12147
21	61	0.6	3337	2	US-08-072-610-1
22	61	0.6	3337	2	US-08-072-610-1
23	61	0.6	3337	3	US-09-092-458-1
24	61	0.6	3337	3	US-08-719-821C-1

25	60.2	0.6	2370	3	US-09-754-947-2	Sequence 2, Appli
26	60	0.6	1956	3	US-08-559-896B-1	Sequence 1, Appli
27	60	0.6	1956	3	US-09-351-794A-1	Sequence 1, Appli
28	59.4	0.6	1141	3	US-09-806-708B-22	Sequence 22, Appli
29	59.4	0.6	4197	2	US-08-682-517-7	Sequence 7, Appli
30	59.4	0.6	4197	2	US-08-682-517-8	Sequence 8, Appli
31	59.4	0.6	14066	3	US-09-601-198-56	Sequence 56, Appli
32	59	0.6	3489	2	US-08-728-323A-1	Sequence 1, Appli
33	59	0.6	3489	3	US-09-298-568-1	Sequence 1, Appli
34	59	0.6	3489	3	US-09-410-359-1	Sequence 1, Appli
35	59	0.6	3489	3	US-09-894-273-1	Sequence 1, Appli
36	59	0.6	32207	2	US-08-770-319-20	Sequence 20, Appli
37	59	0.6	32207	3	US-08-757-669A-20	Sequence 20, Appli
38	59	0.6	32207	3	US-09-230-371A-20	Sequence 20, Appli
39	58.2	0.6	3666	2	US-08-682-517-13	Sequence 13, Appli
40	58.2	0.6	3666	2	US-08-682-517-14	Sequence 14, Appli
41	58.2	0.6	4860	3	US-09-328-352-3221	Sequence 3221, Ap
42	57.4	0.6	979	3	US-09-830-230A-690	Sequence 690, App
43	57.4	0.6	1083	3	US-09-830-230A-689	Sequence 689, App
44	57.4	0.6	1141	3	US-09-806-708B-22	Sequence 22, Appli
45	57.4	0.6	2169	3	US-09-434-408-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1	US-09-841-786-8
Sequence 8, Application US/09841786	
Patent No. 6669940	
GENERAL INFORMATION:	
APPLICANT: NAGARAJ, T. G.	
APPLICANT: STEWART, GEORGE C.	
APPLICANT: NARAYANAN, SANDEEV K.	
APPLICANT: CHENGAPPA, M. M.	
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUROTOXIN	
TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF	
FILE REFERENCE: 30296	
CURRENT APPLICATION NUMBER: US/09/841,786	
CURRENT FILING DATE: 2001-04-24	
PRIOR APPLICATION NUMBER: 09/558,257	
PRIOR FILING DATE: 2000-04-25	
NUMBER OF SEQ ID NOS: 15	
SOFTWARE: Patentin Ver. 2.1	
SEQ ID NO 8	
LENGTH: 9726	
TYPE: DNA	
ORGANISM: Fusobacterium necrophorum	
US-09-841-786-8	
Query Match	100.0%; Score 9726; DB 3; Length 9726;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 9726; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1	ATGACGGGATCAAAATTAAGTTCAGAGGAGAGAGATTCAGATTCTTAAATTTT
1	ATGACGGGATCAAAATTAAGTTCAGAGGAGAGAGATTCAGATTCTTAAATTTT
61	GTTTTATGATTTTGGATGTTGTTGATTAACATTAAGCGGTAGGCTTAATATCAATC
61	GTTTTATGATTTTGGATGTTGTTGATTAACATTAAGCGGTAGGCTTAATATCAATC
61	GTTTTATGATTTTGGATGTTGTTGATTAACATTAAGCGGTAGGCTTAATATCAATC
121	ACCGGAGCTGAGATTTTGGACCAAAATTAAGGATTAATGTTTATGACATTACT
121	ACCGGAGCTGAGATTTTGGACCAAAATTAAGGATTAATGTTTATGACATTACT
121	ACCGGAGCTGAGATTTTGGACCAAAATTAAGGATTAATGTTTATGACATTACT
181	ACCAACAGATTCAGAGGAGAACGCTTTAAGTTTATGATTTTCTTTAAGAGAA
181	ACCAACAGATTCAGAGGAGAACGCTTTAAGTTTATGATTTTCTTTAAGAGAA
241	ATAATATATGCAATCTATATTTTGGGAAAGAAATAGTAGCGGGGTAATATCTTTT
241	ATAATATATGCAATCTATATTTTGGGAAAGAAATAGTAGCGGGGTAATATCTTTT

301 AACCTTGCATGAGAAAAATGAGTAGATGGGATTTATCAACGGAATTCGAGAAAAATAA 360
Db AACCTTGCATGAGAAAAATGAGTAGATGGGATTTATCAACGGAATTCGAGAAAAATAA 360
361 ATTGGAGAGAAATTTATATTTCTTAAAGCTCGAGAGGAGATGGCAGTAGAGAAAAATGAGATT 420
Db ATTGGAGAGAAATTTATATTTCTTAAAGCTCGAGAGGAGATGGCAGTAGAGAAAAATGAGATT 420
421 ATCAATGCTGGTCTTTTCATTCATTAATTCGAAAAACAAGATGATTTTAAAGAGCTTTG 480
Db ATCAATGCTGGTCTTTTCATTCATTAATTCGAAAAACAAGATGATTTTAAAGAGCTTTG 480
421 ATCAATGCTGGTCTTTTCATTCATTAATTCGAAAAACAAGATGATTTTAAAGAGCTTTG 480
481 GAAGAGCCAAACATGATGTAAGTTTTTAAATGGAATCATTCGAGATGAGAAAGTAAAA 540
Db GAAGAGCCAAACATGATGTAAGTTTTTAAATGGAATCATTCGAGATGAGAAAGTAAAA 540
541 ATTCCATTGAATCCGAAATGGAAGCATTAACGGTAGAAGAAAAATCAATGCTGTTGAAGC 600
Db ATTCCATTGAATCCGAAATGGAAGCATTAACGGTAGAAGAAAAATCAATGCTGTTGAAGC 600
601 ATCCGTTTATATGCGCGCGGATTTAGATTGAAAGATCTGCAATCTTAAGACAGAAATT 660
Db ATCCGTTTATATGCGCGCGGATTTAGATTGAAAGATCTGCAATCTTAAGACAGAAATT 660
601 ATCCGTTTATATGCGCGCGGATTTAGATTGAAAGATCTGCAATCTTAAGACAGAAATT 660
661 ACAGATTTTAAAAATTTAGTCAATATTAAGTATCGAATPAAATTTCTGCTGACCGAGAT 720
Db ACAGATTTTAAAAATTTAGTCAATATTAAGTATCGAATPAAATTTCTGCTGACCGAGAT 720
721 TTTAAAGCTACCAAGACAAAATCTGAGATATTAATTTCTTACGCTCAGATAGATTCTCT 780
Db TTTAAAGCTACCAAGACAAAATCTGAGATATTAATTTCTTACGCTCAGATAGATTCTCT 780
721 TTTAAAGCTACCAAGACAAAATCTGAGATATTAATTTCTTACGCTCAGATAGATTCTCT 780
781 CAAAAAGCTATGAGAAAAATTTCACTGTTGAAAGAGATGAGAAATATGTAAAGAA 840
Db CAAAAAGCTATGAGAAAAATTTCACTGTTGAAAGAGATGAGAAATATGTAAAGAA 840
781 CAAAAAGCTATGAGAAAAATTTCACTGTTGAAAGAGATGAGAAATATGTAAAGAA 840
841 AATACCAAGCAAAATATGGAATCTGATGCTGATTTGAAAGCAGATGAGAAATTAATAAAT 900
Db AATACCAAGCAAAATATGGAATCTGATGCTGATTTGAAAGCAGATGAGAAATTAATAAAT 900
841 AATACCAAGCAAAATATGGAATCTGATGCTGATTTGAAAGCAGATGAGAAATTAATAAAT 900
901 AGTGGAAAGCTACAAATGAGAGATTTATAAAGAAAGAGGAAAAAGAACTTATTAAC 960
Db AGTGGAAAGCTACAAATGAGAGATTTATAAAGAAAGAGGAAAAAGAACTTATTAAC 960
901 AGTGGAAAGCTACAAATGAGAGATTTATAAAGAAAGAGGAAAAAGAACTTATTAAC 960
961 ACTCCTTTAAGTTATCAGATGAGAGATTTCCGTAAGATPAAATTAAGAAAGTCTATA 1020
Db ACTCCTTTAAGTTATCAGATGAGAGATTTCCGTAAGATPAAATTAAGAAAGTCTATA 1020
961 ACTCCTTTAAGTTATCAGATGAGAGATTTCCGTAAGATPAAATTAAGAAAGTCTATA 1020
1021 GGAAGAAATGTCAGATTACAGCTGAGCAAGAAATTTCTATGATGCAATTTAGTACT 1080
Db GGAAGAAATGTCAGATTACAGCTGAGCAAGAAATTTCTATGATGCAATTTAGTACT 1080
1021 GGAAGAAATGTCAGATTACAGCTGAGCAAGAAATTTCTATGATGCAATTTAGTACT 1080
1081 AAGCTTGCAGAAAGCTCTTTTACGTTGTTACAGGTTCTATTTCTCTATCAATTTAAT 1140
Db AAGCTTGCAGAAAGCTCTTTTACGTTGTTACAGGTTCTATTTCTCTATCAATTTAAT 1140
1081 AAGCTTGCAGAAAGCTCTTTTACGTTGTTACAGGTTCTATTTCTCTATCAATTTAAT 1140
1141 GGAATTTTAAAGTTATGACAGTAAGTCCAGTTCGTTATTTGAAAGATGCCAAAGTC 1200
Db GGAATTTTAAAGTTATGACAGTAAGTCCAGTTCGTTATTTGAAAGATGCCAAAGTC 1200
1141 GGAATTTTAAAGTTATGACAGTAAGTCCAGTTCGTTATTTGAAAGATGCCAAAGTC 1200
1201 GAAGCAACAGAGAGAGAAAGCAATATTCATCTTCAAGTGAAGTAAGCAATTTGGA 1260
Db GAAGCAACAGAGAGAGAAAGCAATATTCATCTTCAAGTGAAGTAAGCAATTTGGA 1260
1201 GAAGCAACAGAGAGAGAAAGCAATATTCATCTTCAAGTGAAGTAAGCAATTTGGA 1260
1261 GCAGCTACTTCTCAATTAATAATTTCAATTTATTTGAGAGAAAGCCATGAGAAATCTT 1320
Db GCAGCTACTTCTCAATTTAAATTTCAATTTATTTGAGAGAAAGCCATGAGAAATCTT 1320
1261 GCAGCTACTTCTCAATTTAAATTTCAATTTATTTGAGAGAAAGCCATGAGAAATCTT 1320
1321 CTCAGTATTCGAGAGCGGATATTTCTGCAAAAAAGTAATTCGATGTAATTTGAAGA 1380
Db CTCAGTATTCGAGAGCGGATATTTCTGCAAAAAAGTAATTCGATGTAATTTGAAGA 1380
1321 CTCAGTATTCGAGAGCGGATATTTCTGCAAAAAAGTAATTCGATGTAATTTGAAGA 1380

1381 GAAGTAAATCGAAGGAGAGAGAGATATTACTTCAAAATCTGAAAAATACTATTGATGCT 1440
Db GAAGTAAATCGAAGGAGAGAGAGATATTACTTCAAAATCTGAAAAATACTATTGATGCT 1440
1381 GAAGTAAATCGAAGGAGAGAGAGATATTACTTCAAAATCTGAAAAATACTATTGATGCT 1440
1441 TCTGTTTCTGTTGAACGATGAGAGATTTCCATTAAGTACGCTTTTCAGTATTTGAGAG 1500
Db TCTGTTTCTGTTGAACGATGAGAGATTTCCATTAAGTACGCTTTTCAGTATTTGAGAG 1500
1441 TCTGTTTCTGTTGAACGATGAGAGATTTCCATTAAGTACGCTTTTCAGTATTTGAGAG 1500
1501 GAAAGAGAAAAATTAATCTTCCGTCAAGATTGCTAAAGAGCAAAAGTAAATTCAGAAAACG 1560
Db GAAAGAGAAAAATTAATCTTCCGTCAAGATTGCTAAAGAGCAAAAGTAAATTCAGAAAACG 1560
1501 GAAAGAGAAAAATTAATCTTCCGTCAAGATTGCTAAAGAGCAAAAGTAAATTCAGAAAACG 1560
1561 GATGATGTAATGAGAGAGAGAGATTAATTCATTCGAGCTGCTGTAAAGAGTGA 1620
Db GATGATGTAATGAGAGAGAGAGATTAATTCATTCGAGCTGCTGTAAAGAGTGA 1620
1561 GATGATGTAATGAGAGAGAGAGATTAATTCATTCGAGCTGCTGTAAAGAGTGA 1620
1621 TTGGGGGATAGTGGTAAATGGGGTTGTGGCTGCAAAATTTTCTAATCTAATGCTTCCCTCC 1680
Db TTGGGGGATAGTGGTAAATGGGGTTGTGGCTGCAAAATTTTCTAATCTAATGCTTCCCTCC 1680
1621 TTGGGGGATAGTGGTAAATGGGGTTGTGGCTGCAAAATTTTCTAATCTAATGCTTCCCTCC 1680
1681 CGTATAGATGATGATGATATCTACATGCCAAGAGAGCACTAAATGTCAGAGCTCATTAAC 1740
Db CGTATAGATGATGATGATATCTACATGCCAAGAGAGCACTAAATGTCAGAGCTCATTAAC 1740
1681 CGTATAGATGATGATGATATCTACATGCCAAGAGAGCACTAAATGTCAGAGCTCATTAAC 1740
1741 ATTACTAAAAATAGTGTCTTCCGCAAAACAGAGATCTGATTTGGGAACTTCCAACTTATGAAT 1800
Db ATTACTAAAAATAGTGTCTTCCGCAAAACAGAGATCTGATTTGGGAACTTCCAACTTATGAAT 1800
1741 ATTACTAAAAATAGTGTCTTCCGCAAAACAGAGATCTGATTTGGGAACTTCCAACTTATGAAT 1800
1801 GATCACTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db GATCACTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
1801 GATCACTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
1861 GGAAGAGACAGTGTCAATGAGAAATTAAGAAATTAAGTAAAGAACTTATTAAGTGGCT 1920
Db GGAAGAGACAGTGTCAATGAGAAATTAAGAAATTAAGTAAAGAACTTATTAAGTGGCT 1920
1861 GGAAGAGACAGTGTCAATGAGAAATTAAGAAATTAAGTAAAGAACTTATTAAGTGGCT 1920
1921 GTGCTGCAACCATAGCAAAATCATATAATTTCTGCTTCTGTCGCAATAGAGAGAGTGA 1980
Db GTGCTGCAACCATAGCAAAATCATATAATTTCTGCTTCTGTCGCAATAGAGAGAGTGA 1980
1921 GTGCTGCAACCATAGCAAAATCATATAATTTCTGCTTCTGTCGCAATAGAGAGAGTGA 1980
1921 GTGCTGCAACCATAGCAAAATCATATAATTTCTGCTTCTGTCGCAATAGAGAGAGTGA 1980
1981 AGACTTTCTTCAAGAGTGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db AGACTTTCTTCAAGAGTGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
1981 AGACTTTCTTCAAGAGTGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
2041 GGAAGAGACTACGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2100
Db GGAAGAGACTACGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2100
2041 GGAAGAGACTACGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2100
2101 GGAATGCAAGAGTTTTATGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
Db GGAATGCAAGAGTTTTATGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
2101 GGAATGCAAGAGTTTTATGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
2161 CATGCTGAATGATGATGAT 2220
Db CATGCTGAATGATGATGAT 2220
2161 CATGCTGAATGATGATGAT 2220
2221 AAAAATCTTCAAAAATGAGCAAAAGTCTGTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
Db AAAAATCTTCAAAAATGAGCAAAAGTCTGTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
2221 AAAAATCTTCAAAAATGAGCAAAAGTCTGTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
2281 TTTGAT 2340
Db TTTGAT 2340
2281 TTTGAT 2340
2341 TTATTTGAATGATTTTCAAGAAAAATTTGATGAGAAAAACCGAGCTTTTACTAAATGAGTGA 2400
Db TTATTTGAATGATTTTCAAGAAAAATTTGATGAGAAAAACCGAGCTTTTACTAAATGAGTGA 2400
2341 TTATTTGAATGATTTTCAAGAAAAATTTGATGAGAAAAACCGAGCTTTTACTAAATGAGTGA 2400
2401 AGAATGACAAATTAATTTCTCCGATGAGAACTTCAAAAACAGGAATCTGCTATGAGAAATTTGCA 2460
Db AGAATGACAAATTAATTTCTCCGATGAGAACTTCAAAAACAGGAATCTGCTATGAGAAATTTGCA 2460
2401 AGAATGACAAATTAATTTCTCCGATGAGAACTTCAAAAACAGGAATCTGCTATGAGAAATTTGCA 2460
2461 AACTATGTTCAAGGAGAAATGAAAAAATTAAGAGAAAAATTAACGAAAGAGATTTAAAGCT 2520

Db 2461 AACTATGTCAGCGGAAATGAAAATTAGAGAAAATTACGAAAGATTTTAAAGCT 2520
Qy 2521 TTTTCAGAGAGATTAGTGAATGATTAAGAACTTTGATTTTACAGAGATAGAAAT 2580
Db 2521 TTTTCAGAGAGATTAGTGAATGATTAAGAACTTTGATTTTACAGAGATAGAAAT 2580
Qy 2581 TATGCAATTTTACACTTTTACCTTCCGAGCTAATGAGAAAGAGATGTTCTCT 2640
Db 2581 TATGCAATTTTACACTTTTACCTTCCGAGCTAATGAGAAAGAGATGTTCTCT 2640
Qy 2641 GTGGAGAGAGCTGTTTCGTGGGTAGACAGAGAAATTATAGCAAGTATCCGTTGAAAA 2700
Db 2641 GTGGAGAGAGCTGTTTCGTGGGTAGACAGAGAAATTATAGCAAGTATCCGTTGAAAA 2700
Qy 2701 GGAGCTAAACTTGCTGCAAAAAAAGATTAAATATATAAGCTATGCAATTAAGCAGAAACA 2760
Db 2701 GGAGCTAAACTTGCTGCAAAAAAAGATTAAATATATAAGCTATGCAATTAAGCAGAAACA 2760
Qy 2761 GTGAATTTTGTGGAATATTGGAATTTGGAAGAGAGATCAATCCGGAAGTGCAGTCGGA 2820
Db 2761 GTGAATTTTGTGGAATATTGGAATTTGGAAGAGAGATCAATCCGGAAGTGCAGTCGGA 2820
Qy 2821 GGAAGATTAAATGTTCAAGATCGAAAAATTCAGCTATGTAGAAAGCTAAAGAAAAAGCT 2880
Db 2821 GGAAGATTAAATGTTCAAGATCGAAAAATTCAGCTATGTAGAAAGCTAAAGAAAAAGCT 2880
Qy 2881 GAATTAATCAGAGAGAAAAATTATATGAGATGCAATGGAACAGACTTTTTCATGTAGCGGGA 2940
Db 2881 GAATTAATCAGAGAGAAAAATTATATGAGATGCAATGGAACAGACTTTTTCATGTAGCGGGA 2940
Qy 2941 TCTTTTAAATGTGCTCAGGTGGGAATGCAATCAATGGAATGGAATGAGTGAAGCT 3000
Db 2941 TCTTTTAAATGTGCTCAGGTGGGAATGCAATCAATGGAATGGAATGAGTGAAGCT 3000
Qy 3001 ATCAGTAAGGCAAGATTTTCATGATGAGCAAGCAATTTTGAAGCTAATTAATAAAAAAT 3060
Db 3001 ATCAGTAAGGCAAGATTTTCATGATGAGCAAGCAATTTTGAAGCTAATTAATAAAAAAT 3060
Qy 3061 GCTTTTAAACAGTAAGATATACCTCTGTTTGGAAATGCTCCGGTCAACGGGAATCGGA 3120
Db 3061 GCTTTTAAACAGTAAGATATACCTCTGTTTGGAAATGCTCCGGTCAACGGGAATCGGA 3120
Qy 3121 ACGAAAAATGCGCGGTCGCGGTTGCTGTTGCGTAATGATTAATTCACAAAGAA 3180
Db 3121 ACGAAAAATGCGCGGTCGCGGTTGCTGTTGCGTAATGATTAATTCACAAAGAA 3180
Qy 3181 GCTTCATTTGAGATATGACGAAGAGCAAGATTAATGATTAAGATTAAGATGATGA 3240
Db 3181 GCTTCATTTGAGATATGACGAAGAGCAAGATTAATGATTAAGATTAAGATGATGA 3240
Qy 3241 GTTAACAGTAATGCGGAATCTTTAGAGATGATGCAAAAAAGCAACGGAATCAACAGT 3300
Db 3241 GTTAACAGTAATGCGGAATCTTTAGAGATGATGCAAAAAAGCAACGGAATCAACAGT 3300
Qy 3301 ATTTCTGTTGCGCGAGAAATTAAAGTTGGAAGTAAACGAGTGAAGAAAAACCGAAA 3360
Db 3301 ATTTCTGTTGCGCGAGAAATTAAAGTTGGAAGTAAACGAGTGAAGAAAAACCGAAA 3360
Qy 3361 TCGAAGAAAGACCAAGAGGATTTTGTGCAAAATCGGAACAATGGAATCTCTGTAAGA 3420
Db 3361 TCGAAGAAAGACCAAGAGGATTTTGTGCAAAATCGGAACAATGGAATCTCTGTAAGA 3420
Qy 3421 AATTAATTAATACGATAGTATGATTCATTAACAGAAAAATTAACAATTAATTCATTCGA 3480
Db 3421 AATTAATTAATACGATAGTATGATTCATTAACAGAAAAATTAACAATTAATTCATTCGA 3480
Qy 3481 GGAATTAATAAAGCGGGAATCTTCTTCGACGTTTCTCAATCTCCGATTAAGAACCG 3540
Db 3481 GGAATTAATAAAGCGGGAATCTTCTTCGACGTTTCTCAATCTCCGATTAAGAACCG 3540
Qy 3541 TCTTTCAATTTGGAGCTTCTGGAAGTGTCTTTCAATTAATTAATAAGAAACATCT 3600
Db 3541 TCTTTCAATTTGGAGCTTCTGGAAGTGTCTTTCAATTAATTAATAAGAAACATCT 3600

Db 3541 TCTTTCAATTTGGAGCTTCTGGAAGTGTCTTTCAATTAATTAATAAGAAACATCT 3600
Qy 3601 GCTGTCGTAATGAGATTAATAATTTGAAAGGAGCAAAATTAAGAGTGAAGTGA 3660
Db 3601 GCTGTCGTAATGAGATTAATAATTTGAAAGGAGCAAAATTAAGAGTGAAGTGA 3660
Qy 3661 TCTTCTGATTTCACTTTTGTGAGACATGGGGCGGATCTGTCACCTTCACTGGAATCAT 3720
Db 3661 TCTTCTGATTTCACTTTTGTGAGACATGGGGCGGATCTGTCACCTTCACTGGAATCAT 3720
Qy 3721 ATTGGAAGTGAATAATGCAACATGATGCTGTTTACGTCGAGCGGCTGTTAAATAT 3780
Db 3721 ATTGGAAGTGAATAATGCAACATGATGCTGTTTACGTCGAGCGGCTGTTAAATAT 3780
Qy 3781 ATTCAAATTAATAAGAGTCTTTGTTTAAATTAATGATGATTTTCCGAAAGCCAAATTAAT 3840
Db 3781 ATTCAAATTAATAAGAGTCTTTGTTTAAATTAATGATGATTTTCCGAAAGCCAAATTAAT 3840
Qy 3841 AAAGTAATGCTTTGAGTGAAGAGAACTCAAGTACAGAGAGAGAGAGTGGAGACAGTT 3900
Db 3841 AAAGTAATGCTTTGAGTGAAGAGAACTCAAGTACAGAGAGAGAGAGTGGAGACAGTT 3900
Qy 3901 AAAGAAAGTGAAGAGCAAGAAAAAGTTATGTTTGGAACTTCTGTTCTATCAACTTA 3960
Db 3901 AAAGAAAGTGAAGAGCAAGAAAAAGTTATGTTTGGAACTTCTGTTCTATCAACTTA 3960
Qy 3961 GTGAACATGAAGTTTCTGCAAAATCAGAAAAATTAACAGTACGAGAGAAATCTGAAAGC 4020
Db 3961 GTGAACATGAAGTTTCTGCAAAATCAGAAAAATTAACAGTACGAGAGAAATCTGAAAGC 4020
Qy 4021 CAAAAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4080
Db 4021 CAAAAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4080
Qy 4081 AATTTCAAGAGTGAAGATGCAATGAACTGATGAGGAGTACTGATGATGATGATGATGATGAT 4140
Db 4081 AATTTCAAGAGTGAAGATGCAATGAACTGATGAGGAGTACTGATGATGATGATGATGATGAT 4140
Qy 4141 AAACAACAAAGTAAATGCTTTTATTTAGTGGGAGATTAATCAATGATTAATGAGCGGAC 4200
Db 4141 AAACAACAAAGTAAATGCTTTTATTTAGTGGGAGATTAATCAATGATTAATGAGCGGAC 4200
Qy 4201 GCAAAAGCTCTTTTGAACCACTCAAGTACGTCGAGTGAAGAGAGAGAGCAAT 4260
Db 4201 GCAAAAGCTCTTTTGAACCACTCAAGTACGTCGAGTGAAGAGAGAGAGCAAT 4260
Qy 4261 AGTTCTGAGCGGATTAAGAAATTAATCAAGGAGGCTGTTCTGTCATTAATGATGA 4320
Db 4261 AGTTCTGAGCGGATTAAGAAATTAATCAAGGAGGCTGTTCTGTCATTAATGATGA 4320
Qy 4321 GAGTGAAGCTAGCGTGAATTAATCTTCATGAGAGAGCTTAATGAAATCAATGTCAT 4380
Db 4321 GAGTGAAGCTAGCGTGAATTAATCTTCATGAGAGAGCTTAATGAAATCAATGTCAT 4380
Qy 4381 GCCAAAGATGCAAGAAAGCTCTGATCAGCAAAAGAAATATCAGGCTTATCAATATGA 4440
Db 4381 GCCAAAGATGCAAGAAAGCTCTGATCAGCAAAAGAAATATCAGGCTTATCAATATGA 4440
Qy 4441 AAAGATTAATAATATTAGAAATGATGATGATTAATCAATGAAATGATTAATGATGATGAT 4500
Db 4441 AAAGATTAATAATATTAGAAATGATGATGATTAATCAATGAAATGATTAATGATGATGAT 4500
Qy 4501 AAAGAACTAGAGAAAGCAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
Db 4501 AAAGAACTAGAGAAAGCAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
Qy 4561 TCGGTTCTGGAACCGATTAATCCGCTGGAAGAGTATGATGATGATGATGATGATGATGAT 4620
Db 4561 TCGGTTCTGGAACCGATTAATCCGCTGGAAGAGTATGATGATGATGATGATGATGATGAT 4620
Qy 4621 AATTAATTAATAAGCAATTTGATGAGAGCAATTAAGAAAGAGAGAGAGAGAGAGAGAGAG 4680
Db 4621 AATTAATTAATAAGCAATTTGATGAGAGCAATTAAGAAAGAGAGAGAGAGAGAGAGAGAG 4680

QY 4681 GCGAAACATGTAAATGTGAGCGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGGA 4740
DB 4681 GCGAAACATGTAAATGTGAGCGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGGA 4740
QY 4741 CTGTCTATCAGCAAAAGATGCTTTTTCAGGAATGGGATCTGAGCATGCGAAGACTTATCA 4800
DB 4741 CTGTCTATCAGCAAAAGATGCTTTTTCAGGAATGGGATCTGAGCATGCGAAGACTTATCA 4800
QY 4801 AATGACAGATTGCAAGGTGGAATAAGGAAGAAATTTCTGCGATTCCTTAAATGTGAC 4860
DB 4801 AATGACAGATTGCAAGGTGGAATAAGGAAGAAATTTCTGCGATTCCTTAAATGTGAC 4860
QY 4861 GCAAAATTAATTCATTTCTGCGGTGATGTTGCGGAGCAATGCGGTTCTTCTTACG 4920
DB 4861 GCAAAATTAATTCATTTCTGCGGTGATGTTGCGGAGCAATGCGGTTCTTCTTACG 4920
QY 4921 GCGGTAGAGCTGCTTTTTCGCAATTAATCTCTTCAATAATAAACCCTGCTTGAATTACA 4980
DB 4921 GCGGTAGAGCTGCTTTTTCGCAATTAATCTCTTCAATAATAAACCCTGCTTGAATTACA 4980
QY 4981 GGAACGAAGTAAATCCTTTTAGTGAAGAAATACAAAGTCAATGTAACAAGCTTGAAT 5040
DB 4981 GGAACGAAGTAAATCCTTTTAGTGAAGAAATACAAAGTCAATGTAACAAGCTTGAAT 5040
QY 5041 GATTTCTATTTACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAGCAGGCTGGAATC 5100
DB 5041 GATTTCTATTTACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAGCAGGCTGGAATC 5100
QY 5101 GAGAGAAATGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5160
DB 5101 GAGAGAAATGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5160
QY 5161 GAGTTGAAGAGTAAATCTTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5220
DB 5161 GAGTTGAAGAGTAAATCTTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5220
QY 5221 ATTGCGGAAATGCAAAATGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACA 5280
DB 5221 ATTGCGGAAATGCAAAATGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACA 5280
QY 5281 AATATTGGAACCAATCAGTTATAGCTATTGTAATAAAGCAATGTAATAAATTAACAACGCGGAT 5340
DB 5281 AATATTGGAACCAATCAGTTATAGCTATTGTAATAAAGCAATGTAATAAATTAACAACGCGGAT 5340
QY 5341 GATCAAGATGAGAAATAATATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5400
DB 5341 GATCAAGATGAGAAATAATATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5400
QY 5401 GCAGTCGAGTTGAGAGCAAAAGAGCTCTGTCAGCAAGAGCTTCTGCAAGTACTAC 5460
DB 5401 GCAGTCGAGTTGAGAGCAAAAGAGCTCTGTCAGCAAGAGCTTCTGCAAGTACTAC 5460
QY 5461 TTGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAAGGAA 5520
DB 5461 TTGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAAGGAA 5520
QY 5521 GAAATTAATGAGAAATAAGCAAAAGCAATGTTAATGTTCTAGCTGAAATTCGAGTGA 5580
DB 5521 GAAATTAATGAGAAATAAGCAAAAGCAATGTTAATGTTCTAGCTGAAATTCGAGTGA 5580
QY 5581 GTGTCACAAATGCGACAGTGTCTTCCGAGCAATGAGCAAGCTGCGATGAGACTGGA 5640
DB 5581 GTGTCACAAATGCGACAGTGTCTTCCGAGCAATGAGCAAGCTGCGATGAGACTGGA 5640
QY 5641 GTAGCAGTTAATTAATTAACAAATACTTCTGCACATATAAATAATAGTCTCAAAAT 5700
DB 5641 GTAGCAGTTAATTAATTAACAAATACTTCTGCACATATAAATAATAGTCTCAAAAT 5700
QY 5701 GTACGAATATGCTTGTGTAAAGCAAAATCTCATCTATTAATAAATGGAATTGGA 5760
DB 5701 GTACGAATATGCTTGTGTAAAGCAAAATCTCATCTATTAATAAATGGAATTGGA 5760

QY 5761 GCTGAGATTGAGAGCTGAGAGAGCTGAGATGACAGGTTCTGTAGCAGTGAATTAAGATTGTA 5820
DB 5761 GCTGAGATTGAGAGCTGAGAGAGCTGAGATGACAGGTTCTGTAGCAGTGAATTAAGATTGTA 5820
QY 5821 AATAATACGATAGCAAAATTAATATCTGCAAAATTCACCTGGAAGGCAAAATGTGGAATT 5880
DB 5821 AATAATACGATAGCAAAATTAATATCTGCAAAATTCACCTGGAAGGCAAAATGTGGAATT 5880
QY 5881 ATTACAGAGCTGATGAGGCTTAAATGCTTAATATGAGGAAACAGTGTGAGTGGCCCGT 5940
DB 5881 ATTACAGAGCTGATGAGGCTTAAATGCTTAATATGAGGAAACAGTGTGAGTGGCCCGT 5940
QY 5941 GCAGCAATAGAGGCTCAACAGGTGATGAATGAATTAACAGATCTACAAAGCATATGTA 6000
DB 5941 GCAGCAATAGAGGCTCAACAGGTGATGAATGAATTAACAGATCTACAAAGCATATGTA 6000
QY 6001 AAAATTTCTACAGTATGCTTTAAAGAAAGAAACAGATGATTAATTAATCTACCAAGGCA 6060
DB 6001 AAAATTTCTACAGTATGCTTTAAAGAAAGAAACAGATGATTAATTAATCTACCAAGGCA 6060
QY 6061 GTAGTAAAGTGTAGATTAAGTATTCAAAATCTTAATATTAAGAAAGCTTATACAA 6120
DB 6061 GTAGTAAAGTGTAGATTAAGTATTCAAAATCTTAATATTAAGAAAGCTTATACAA 6120
QY 6121 AAAAGAAAAATAGTAATAAAGAAAGAAATTTGTTACCAATAGTCACTACTATCTTTA 6180
DB 6121 AAAAGAAAAATAGTAATAAAGAAAGAAATTTGTTACCAATAGTCACTACTATCTTTA 6180
QY 6181 AAATCTTTATTTGCAAAATCCGCTGTGTCAGAGCAACCCGAGTGGCAGAACTGTTAAT 6240
DB 6181 AAATCTTTATTTGCAAAATCCGCTGTGTCAGAGCAACCCGAGTGGCAGAACTGTTAAT 6240
QY 6241 ATCAACAAAGTTTATGAGAAACAGAAAGCTCTTGTAGAAATTTCTATTAATAGCAAAA 6300
DB 6241 ATCAACAAAGTTTATGAGAAACAGAAAGCTCTTGTAGAAATTTCTATTAATAGCAAAA 6300
QY 6301 CATTATTTCTGTAAATCAGAGATTAACAGAAATTCATTCGAGTATGAGTTCTGTGGT 6360
DB 6301 CATTATTTCTGTAAATCAGAGATTAACAGAAATTCATTCGAGTATGAGTTCTGTGGT 6360
QY 6361 GTTGTGGAATGTAAGAGTATGAGAGCTTCTTGTATACCAATATTAATAAAGAAATACC 6420
DB 6361 GTTGTGGAATGTAAGAGTATGAGAGCTTCTTGTATACCAATATTAATAAAGAAATACC 6420
QY 6421 AAGAAGAGTTGGAAGAAACATCAATGTCGATGAGGTTTCGAGAAAGAGTGAATTT 6480
DB 6421 AAGAAGAGTTGGAAGAAACATCAATGTCGATGAGGTTTCGAGAAAGAGTGAATTT 6480
QY 6481 ACACAGATTTCTAAGCAAGAAATTTCTTTTGAAGTCGAGATCGACAGCGGGTA 6540
DB 6481 ACACAGATTTCTAAGCAAGAAATTTCTTTTGAAGTCGAGATCGACAGCGGGTA 6540
QY 6541 GGAACCGGAGTGGCAGGAAACCGTTTCCGTAAATCAATTTGCAAGAAAGACGAAATGAT 6600
DB 6541 GGAACCGGAGTGGCAGGAAACCGTTTCCGTAAATCAATTTGCAAGAAAGACGAAATGAT 6600
QY 6601 GTGGAAGAGCAAGATTTTGTATAAAGAAACCTAGATTAACAGCAAAAGCTTATAGTCT 6660
DB 6601 GTGGAAGAGCAAGATTTTGTATAAAGAAACCTAGATTAACAGCAAAAGCTTATAGTCT 6660
QY 6661 GTTCAATTTGGAATGCGCAGTGTGAGTGGCTGCAAAAGAGCTGGAATTTGAGACACA 6720
DB 6661 GTTCAATTTGGAATGCGCAGTGTGAGTGGCTGCAAAAGAGCTGGAATTTGAGACACA 6720
QY 6721 GTGCGAGTTACCAAGATGATTAACAACGAGAGCAGAGTGAATAATTTCTTAATAATATG 6780
DB 6721 GTGCGAGTTACCAAGATGATTAACAACGAGAGCAGAGTGAATAATTTCTTAATAATATG 6780
QY 6781 ACTGGAACCAAGTATGATGATTAATGCAAGAAATGATTAATTAATCAAGGATCTGGAATCGGT 6840
DB 6781 ACTGGAACCAAGTATGATGATTAATGCAAGAAATGATTAATTAATCAAGGATCTGGAATCGGT 6840
QY 6841 TCAGCCGAGCTGGAATTTCTTGACGCCGAGTATCTGAGTGTCTTCTGCAATAATATTT 6900

Db 6841 TCAAGCCGAGAGCTGAGATTCTTGACACCGAGATCTGAGAGTGGTTCTGTCAATAATAT 6900
Qy 6901 GCAAAATTAAGTAAGAAACAGATATGCATCATAGTACTTTACACTCTTTCTATCATGATGTAAT 6960
Db 6901 GCAAAATTAAGTAAGAAACAGATATGCATCATAGTACTTTACACTCTTTCTATCATGATGTAAT 6960
Qy 6961 GTAAAAAGCTCTTAATTAATTTTCGAATTCCTTGACAGCCGGTGGAGAGAGCCCGAGTCTT 7020
Db 6961 GTAAAAAGCTCTTAATTAATTTTCGAATTCCTTGACAGCCGGTGGAGAGAGCCCGAGTCTT 7020
Qy 7021 GGAGAGCTTACCGGAGTGGTTCTGTGAACATTAATAGTTAGTTAGTACTGAGAT 7080
Db 7021 GGAGAGCTTACCGGAGTGGTTCTGTGAACATTAATAGTTAGTTAGTACTGAGAT 7080
Qy 7081 CACAATTAATCTGATTTGACTTCCTGACAGAAAAAGTAATGTAAACGACAAAGAGAA 7140
Db 7081 CACAATTAATCTGATTTGACTTCCTGACAGAAAAAGTAATGTAAACGACAAAGAGAA 7140
Qy 7141 AAAAATATTAAAGCAAAACAGCAAGATGCAAGATTCGAGAGACGAATCGAGCCAT 7200
Db 7141 AAAAATATTAAAGCAAAACAGCAAGATGCAAGATTCGAGAGACGAATCGAGCCAT 7200
Qy 7201 GCTTGCTTAATTAATTTTGGAAACGCTGTAGAAAGATAGAAAAATTCGAGAGAAAGGA 7260
Db 7201 GCTTGCTTAATTAATTTTGGAAACGCTGTAGAAAGATAGAAAAATTCGAGAGAAAGGA 7260
Qy 7261 ACAGAAAGTTTAAAACTTTAGACGAAGTTAAACAAAGAACAGATTAATAATGAT 7320
Db 7261 ACAGAAAGTTTAAAACTTTAGACGAAGTTAAACAAAGAACAGATTAATAATGAT 7320
Qy 7321 GCTACGAAAAAAATCTTACAAATCAAGAGTATTTCTACAGAAATCTTCTGTAAAAAGC 7380
Db 7321 GCTACGAAAAAAATCTTACAAATCAAGAGTATTTCTACAGAAATCTTCTGTAAAAAGC 7380
Qy 7381 GATAGAGAGATATCTCAAGAGAAAGAAATTAAGCCATTGTGAAGCTTCTGATATAT 7440
Db 7381 GATAGAGAGATATCTCAAGAGAAAGAAATTAAGCCATTGTGAAGCTTCTGATATAT 7440
Qy 7441 GGAATAAATATGATATATTAACAACAGAGCAAGAAATATATATCACTTCTAGTGGTGG 7500
Db 7441 GGAATAAATATGATATATTAACAACAGAGCAAGAAATATATATCACTTCTAGTGGTGG 7500
Qy 7501 GGAATCTGACAGTCTTGCTTCCGATCAGAAACGTGGCAGTTAACAAATTTAAAAAGAAAT 7560
Db 7501 GGAATCTGACAGTCTTGCTTCCGATCAGAAACGTGGCAGTTAACAAATTTAAAAAGAAAT 7560
Qy 7561 TCCGAGATTAATCTTTGTAATTTCTTTGTAAGAGAGCTGAAAAAGTAAATGTTAGATCG 7620
Db 7561 TCCGAGATTAATCTTTGTAATTTCTTTGTAAGAGAGCTGAAAAAGTAAATGTTAGATCG 7620
Qy 7621 GATATTAACAGAAATGTTGCTTTAACAGCATCAAGGTCCTGAGAGCATTTGGGAATA 7680
Db 7621 GATATTAACAGAAATGTTGCTTTAACAGCATCAAGGTCCTGAGAGCATTTGGGAATA 7680
Qy 7681 GGAAGCTGCTATGCAAGATTAATTAATCTAATGGAAGATCAAAATCTGATTAATAATTTCT 7740
Db 7681 GGAAGCTGCTATGCAAGATTAATTAATCTAATGGAAGATCAAAATCTGATTAATAATTTCT 7740
Qy 7741 AAGCTATTAGAAAAAATATGATGTTATTTGTAAGAGATTAATCGAAATTTGAGAGCGGA 7800
Db 7741 AAGCTATTAGAAAAAATATGATGTTATTTGTAAGAGATTAATCGAAATTTGAGAGCGGA 7800
Qy 7801 GCAAAAGAGTTAACGTAAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAAGCAAGAAAT 7860
Db 7801 GCAAAAGAGTTAACGTAAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAAGCAAGAAAT 7860
Qy 7861 GAATGAATTCAGAGGTTGAAATTTGAGAGAGATATTTCAATGAGAAATATAGATTAAT 7920
Db 7861 GAATGAATTCAGAGGTTGAAATTTGAGAGAGATATTTCAATGAGAAATATAGATTAAT 7920
Qy 7921 AGCCCTTTCAAGAAATTTGAGAGAAATCAATGTCAAATGTGAAAAAGAAAAAGAGATG 7980
Db 7921 AGCCCTTTCAAGAAATTTGAGAGAAATCAATGTCAAATGTGAAAAAGAAAAAGAGATG 7980

Db 7921 AGCCCTTTCAAGAAATTTGAGAGAAATCAATGTCAAATGTGAAAAAGAAAAAGAGATG 7980
Qy 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGAGGGGAGAAATTTATTCGGA 8040
Db 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGAGGGGAGAAATTTATTCGGA 8040
Qy 8041 GCAAAAAGATGCGGAGAGCTCTTAATTTGAAAGTATAGTAAATCCGGAAGAGATTTT 8100
Db 8041 GCAAAAAGATGCGGAGAGCTCTTAATTTGAAAGTATAGTAAATCCGGAAGAGATTTT 8100
Qy 8101 CATGAGATTAATGTGAATATGAAAGCAACATTAATTAAGTAACAGAGTTTCTAAA 8160
Db 8101 CATGAGATTAATGTGAATATGAAAGCAACATTAATTAAGTAACAGAGTTTCTAAA 8160
Qy 8161 GCAGTAACAGGTTCTGTATTTGGAGAGAGTGGAGTCAACCAAGCGAAGCTATCTGCA 8220
Db 8161 GCAGTAACAGGTTCTGTATTTGGAGAGAGTGGAGTCAACCAAGCGAAGCTATCTGCA 8220
Qy 8221 GGTAAATCTATGATGATGAGAGAGAAATTTGTCAGAACAAATCGATTGAATGCA 8280
Db 8221 GGTAAATCTATGATGATGAGAGAGAAATTTGTCAGAACAAATCGATTGAATGCA 8280
Qy 8281 ATTTCTAAAGTAAAGATTTGGATGAGATTAAGTAACTGCTAAATCTTCTGTAGATCA 8340
Db 8281 ATTTCTAAAGTAAAGATTTGGATGAGATTAAGTAACTGCTAAATCTTCTGTAGATCA 8340
Qy 8341 GGAATTTGAGAGAGAAATTCGCGAGCAGAGTGAATCTTCTACAGCAACAAAGTAACT 8400
Db 8341 GGAATTTGAGAGAGAAATTCGCGAGCAGAGTGAATCTTCTACAGCAACAAAGTAACT 8400
Qy 8401 GAAATCCGATGCTGTTTGAAGAGCAATTAAGAAATTAATGATTAACAAAAAAATAT 8460
Db 8401 GAAATCCGATGCTGTTTGAAGAGCAATTAAGAAATTAATGATTAACAAAAAAATAT 8460
Qy 8461 ATTTCAAGAGTCAATGCTCTTGTCTTAAATGATACAAAGAAATGAACGAATATAGATCT 8520
Db 8461 ATTTCAAGAGTCAATGCTCTTGTCTTAAATGATACAAAGAAATGAACGAATATAGATCT 8520
Qy 8521 TTAAGCGGTAAGCCGATGTCATGACACAGAGAACAAACAAAGCTTTACAGATCAACAG 8580
Db 8521 TTAAGCGGTAAGCCGATGTCATGACACAGAGAACAAACAAAGCTTTACAGATCAACAG 8580
Qy 8581 TTAATCTTCAACATGTAATTCAGAGAAACGATCTCAACTGCGCAAAACCTTGTGCT 8640
Db 8581 TTAATCTTCAACATGTAATTCAGAGAAACGATCTCAACTGCGCAAAACCTTGTGCT 8640
Qy 8641 AAAAATGAAAAATTAAGAAATGTAAGAACTGAGAGAGCTTATGCGAGCGGAACA 8700
Db 8641 AAAAATGAAAAATTAAGAAATGTAAGAACTGAGAGAGCTTATGCGAGCGGAACA 8700
Qy 8701 GCAGCCGTTGAAAAATTAACAAGAGTACTACAGAGCATTTGGTTGCAAGAAATTTGGGA 8760
Db 8701 GCAGCCGTTGAAAAATTAACAAGAGTACTACAGAGCATTTGGTTGCAAGAAATTTGGGA 8760
Qy 8761 ATTTGAGATTAATTTGAAACGATTTGCAAGATTAATGATTTGTAAGATCAACGAGAC 8820
Db 8761 ATTTGAGATTAATTTGAAACGATTTGCAAGATTAATGATTTGTAAGATCAACGAGAC 8820
Qy 8821 GGAACCAAGAGAGCTTGTGCGAAAGATGTAATTTCTGTGAAAAATCAATTTCAAGG 8880
Db 8821 GGAACCAAGAGAGCTTGTGCGAAAGATGTAATTTCTGTGAAAAATCAATTTCAAGG 8880
Qy 8881 GAAACAAATCATCTCAATTTGAAATGAAGCAGAAATTTGGAACCGGAAGTAAATGTA 8940
Db 8881 GAAACAAATCATCTCAATTTGAAATGAAGCAGAAATTTGGAACCGGAAGTAAATGTA 8940
Qy 8941 GATGCTTTGAATGAACCTGATGATGATCTCAACAGAAAAAGTGTGCTATGCTGAAT 9000
Db 8941 GATGCTTTGAATGAACCTGATGATGATCTCAACAGAAAAAGTGTGCTATGCTGAAT 9000
Qy 9001 GGTATTTGAAAAATGTTGAATGAATATGATGTTAAGAAAAATGTAAGCCAAAAATCGGA 9060
Db 9001 GGTATTTGAAAAATGTTGAATGAATATGATGTTAAGAAAAATGTAAGCCAAAAATCGGA 9060

QY 9061 AGACATGCTATTGTAGAAACTACTGAAAAACAAGATATCAAGCATTTTACAGAGCAAAA 9120
DB 9061 AGACATGCTATTGTAGAAACTACTGAAAAACAAGATATCAAGCATTTTACAGAGCAAAA 9120
QY 9121 GTTAAATATTTCTTGGAAAAAGAGACGCTGACGCTGCAATATTCGAATGTACACATT 9180
DB 9121 GTTAAATATTTCTTGGAAAAAGAGACGCTGACGCTGCAATATTCGAATGTACACATT 9180
QY 9181 TCCATGAGATGAGATTTAAAAATTTGGCAAGCAGTATGCAATCTTCTCAATTAATAACC 9240
DB 9181 TCCATGAGATGAGATTTAAAAATTTGGCAAGCAGTATGCAATCTTCTCAATTAATAACC 9240
QY 9241 AAAAAATTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9300
DB 9241 AAAAAATTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9300
QY 9301 GGGGCTGCTGAAGCAAGAGCTGACAGAGCCAAAGCGACATTAATGTAAAGATCAAAATA 9360
DB 9301 GGGGCTGCTGAAGCAAGAGCTGACAGAGCCAAAGCGACATTAATGTAAAGATCAAAATA 9360
QY 9361 AATAGAACTAATTAATGTTGATTAGCAGGAAAAATTAACAGAGGAAACATCAATGTA 9420
DB 9361 AATAGAACTAATTAATGTTGATTAGCAGGAAAAATTAACAGAGGAAACATCAATGTA 9420
QY 9421 TATGCGGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9480
DB 9421 TATGCGGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9480
QY 9481 GCCAAAGTCATGCTGCGACGCTGCTGCGCACTGCGCACTATTTGAAAAATTAAGTAAAA 9540
DB 9481 GCCAAAGTCATGCTGCGACGCTGCTGCGCACTGCGCACTATTTGAAAAATTAAGTAAAA 9540
QY 9541 TTTAATTAATGCGATCCGAGAAATTTTAAAAATTAATCTGCAAGATTTGAAAAGGAACTAAT 9600
DB 9541 TTTAATTAATGCGATCCGAGAAATTTTAAAAATTAATCTGCAAGATTTGAAAAGGAACTAAT 9600
QY 9601 AAAAAAAGTCGCTGAGATCTAATCAGGTAGACTGCTATACGATTAATATATCATGCGAT 9660
DB 9601 AAAAAAAGTCGCTGAGATCTAATCAGGTAGACTGCTATACGATTAATATATCATGCGAT 9660
QY 9661 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAGAGAGGAAAAAGGAAA 9720
DB 9661 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAGAGAGGAAAAAGGAAA 9720
QY 9721 AAATGA 9726
DB 9721 AAATGA 9726

RESULT 2
US-09-841-786-15
Sequence 15, Application US/09841786
Patent No. 6669940
GENERAL INFORMATION:
APPLICANT: NAGARAJA, T. G.
APPLICANT: STEWART, GEORGE C.
APPLICANT: NARAYANAN, SANJEEV K.
APPLICANT: CHENGARPA, M. M.
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF
FILE REFERENCE: 30296
CURRENT FILING DATE: 2001-04-24
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 11130
TYPE: DNA
ORGANISM: Fusobacterium necrophorum
US-09-841-786-15

Query Match 100.0%; Score 9726; DB 3; Length 11130;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGGATCAAAAATTAACGTTGAGAGACAAAGAAAGAGATTCAGATTCTAAAAAA 60
DB 1034 ATGAGCGGATCAAAAATTAACGTTGAGAGACAAAGAAAGAGATTCAGATTCTAAAAAA 1093
QY 61 GTTTTAATGATTTTGGGATTTGTTATTAACAATTAACGTTGAGAGGCTTAATGATCAATC 120
DB 1094 GTTTTAATGATTTTGGGATTTGTTATTAACAATTAACGTTGAGAGGCTTAATGATCAATC 1153
QY 121 ACCGCACTGAGAAATTTTGGAAACAAAATTAAGAAAAAAGATTAATGTTATGACATTACT 180
DB 1154 ACCGCACTGAGAAATTTTGGAAACAAAATTAAGAAAAAAGATTAATGTTATGACATTACT 1213
QY 181 ACAAACAAGATTCAAGGAGACGCTTTTAACGTTTAAATTAATTAATGTTGCTTTAAACAGAA 240
DB 1214 ACAAACAAGATTCAAGGAGACGCTTTTAACGTTTAAATTAATTAATGTTGCTTTAAACAGAA 1273
QY 241 AATAATTAAGCAATCTAATTTTGGGAAAAAAGATTAAGGCGGTAAATATCTTTT 300
DB 1274 AATAATTAAGCAATCTAATTTTGGGAAAAAAGATTAAGGCGGTAAATATCTTTT 1333
QY 301 AACTTGTCAATGGAATAATTTGAAGTAGATGGAATTAACAAGAAATTCGAGAAATTAATA 360
DB 1334 AACTTGTCAATGGAATAATTTGAAGTAGATGGAATTAACAAGAAATTCGAGAAATTAATA 1393
QY 361 ATTGAGAAATTTATATTTCTTAAGCTCGAAGGAGATGCGAGTAGAAAAATAGGATT 420
DB 1394 ATTGAGAAATTTATATTTCTTAAGCTCGAAGGAGATGCGAGTAGAAAAATAGGATT 1453
QY 421 ATCAATGCTGCTCTTTTCAATTTCTAATTTTCAAAACAAGATTAATTAAGAGGCTTTG 480
DB 1454 ATCAATGCTGCTCTTTTCAATTTCTAATTTTCAAAACAAGATTAATTAAGAGGCTTTG 1513
QY 481 GAAGAGCAAAACATGTAAGTTTAAATGAAATCAATCCAGTAGATGAAAAATGTAATA 540
DB 1514 GAAGAGCAAAACATGTAAGTTTAAATGAAATCAATCCAGTAGATGAAAAATGTAATA 1573
QY 541 ATTCATTAAGATCCGAATGGAAGCATTAACGTTAGAGAAAAATCAATGCTGTTGAAGC 600
DB 1574 ATTCATTAAGATCCGAATGGAAGCATTAACGTTAGAGAAAAATCAATGCTGTTGAAGC 1633
QY 601 ATCGTTATATGCGGCGGATTAATGAAGATCTGAATCTAATAAGACAGAAAT 660
DB 1634 ATCGTTATATGCGGCGGATTAATGAAGATCTGAATCTAATAAGACAGAAAT 1693
QY 661 ACAGATTTTAAAAATTTAGTCAATATTAAGTGAATTAATTTGCTGCTGACCGAGAT 720
DB 1694 ACAGATTTTAAAAATTTAGTCAATATTAAGTGAATTAATTTGCTGCTGACCGAGAT 1753
QY 721 TTTAAAGCTACCAAGACAAAATCTGAGATATTAATCTTTCAGCTCACATAGATCTCTCT 780
DB 1754 TTTAAAGCTACCAAGACAAAATCTGAGATATTAATCTTTCAGCTCACATAGATCTCTCT 1813
QY 781 CAAAAGCTAAGGAAAAAATTTCAACTGTTGAAAAGAAATGAAGAAATATGTAAGAA 840
DB 1814 CAAAAGCTAAGGAAAAAATTTCAACTGTTGAAAAGAAATGAAGAAATATGTAAGAA 1873
QY 841 AATACCAAGCAATATTTGAATCTGATGCTGATTTGGAAGCAGATGGAATATTAATAATTT 900
DB 1874 AATACCAAGCAATATTTGAATCTGATGCTGATTTGGAAGCAGATGGAATATTAATAATTT 1933
QY 901 AGTGGAAAGCTACAAATGAGAGATTTAATAAGAAAGGAGGAAAAAAGAACTTATTAAC 960
DB 1934 AGTGGAAAGCTACAAATGAGAGATTTAATAAGAAAGGAGGAAAAAAGAACTTATTAAC 1993
QY 961 ACTCCTTAAGTTATCAGATGTTGAAGCTTCGTTAAGATTAATTAAGAAAGAAAGTCTATA 1020
DB 1994 ACTCCTTAAGTTATCAGATGTTGAAGCTTCGTTAAGATTAATTAAGAAAGAAAGTCTATA 2053

QY 1021 GGAAAGAAATGTTGACATTACAGCTGAAAGCAAAAGATTTCTATGATGCAACTTTAGTTACT 1080
Db 2054 GGAAAGAAATGTTGACATTACAGCTGAAAGCAAAAGATTTCTATGATGCAACTTTAGTTACT 2113
QY 1081 AAGCTTGGAAAGCACTCTTTTGTAGCTTTTGTACAGGTTCTATTTCTCTATCAATTTAAT 1140
Db 2114 AAGCTTGGAAAGCACTCTTTTGTAGCTTTTGTACAGGTTCTATTTCTCTATCAATTTAAT 2173
QY 1141 GGATTTTGTAGTTTATGACAAATAGTTCAGTGTCTGTTATGTAAGAAAGATGCCAAAGTC 1200
Db 2174 GGATTTTGTAGTTTATGACAAATAGTTCAGTGTCTGTTATGTAAGAAAGATGCCAAAGTC 2233
QY 1201 GAAGCAACAAAGAAAGCAAAATATCTCTTTCATGAGGTAAGCACTATGCGA 1260
Db 2234 GAAGCAACAAAGAAAGCAAAATATCTCTTTCATGAGGTAAGCACTATGCGA 2293
QY 1261 GAAGCTACTCTTCCATTAAATAATTCATTTATATTTTGGAAAGCCAAATGGAATCTT 1320
Db 2294 GCAAGCTTACTCTTCCATTAAATAATTCATTTATTTGAGAAAGCCAAATGGAATCTT 2353
QY 1321 CTCAGTATCGAGCGGAGATATATTTCTGCAAAAAGTAAATTCATGTAAGGA 1380
Db 2354 CTCAGTATCGAGCGGAGATATATTTCTGCAAAAAGTAAATTCATGTAAGGA 2413
QY 1381 GAAGTAAATCGAAGGGAAGAGACATATTAATTCGAAATCTGAAATATCTATGATGCT 1440
Db 2414 GAAGTAAATCGAAGGGAAGAGACATATTAATTCGAAATCTGAAATATCTATGATGCT 2473
QY 1441 TCTGTTTCTGTTGGAACGATGAGAGATTCATTAATAAGTAGCTCTTTCAGTATTTGGTAGC 1500
Db 2474 TCTGTTTCTGTTGGAACGATGAGAGATTCATTAATAAGTAGCTCTTTCAGTATTTGGTAGC 2533
QY 1501 GAAGGAAATAATTAATCTTCGTCAGATTTGCTAAAGGAGCAAAAGTAGAATCAGAAAGC 1560
Db 2534 GAAGGAAATAATTAATCTTCGTCAGATTTGCTAAAGGAGCAAAAGTAGAATCAGAAAGC 2593
QY 1561 GATGATGTAATGTGGAAGTGAAGCGATTAATTCATTCGAGCTGCTGTAAAGGTGGA 1620
Db 2594 GATGATGTAATGTGGAAGTGAAGCGATTAATTCATTCGAGCTGCTGTAAAGGTGGA 2653
QY 1621 TTGGGGGATAGTGTGTAATGGGGTGTGGCTGCAAAATATTTCTAACTATAATGCTTCTCC 1680
Db 2654 TTGGGGGATAGTGTGTAATGGGGTGTGGCTGCAAAATATTTCTAACTATAATGCTTCTCC 2713
QY 1681 CCGTATGATGATGATGATATCTACATGCGCAAGAGCGCATTAATGTGAGGCTCATTAAC 1740
Db 2714 CCGTATGATGATGATGATATCTACATGCGCAAGAGCGCATTAATGTGAGGCTCATTAAC 2773
QY 1741 ATTACTAAAAATAGTGTCTGCAACAGAGATGATTTGGGAACTTCCAAATTTATGAAT 1800
Db 2774 ATTACTAAAAATAGTGTCTGCAACAGAGATGATTTGGGAACTTCCAAATTTATGAAT 2833
QY 1801 GATCAGCTTATGCAATCAGGTCTCTAAATTCATTTTGTAGCAATTAACAGCGGTTT 1860
Db 2834 GATCAGCTTATGCAATCAGGTCTCTAAATTCATTTTGTAGCAATTAACAGCGGTTT 2893
QY 1861 GGAAGAGACAGTGTCAATGAGAAATTAAGATTAAGCTAAGCACTTATTTAGTGTGCT 1920
Db 2894 GGAAGAGACAGTGTCAATGAGAAATTAAGATTAAGCTAAGCACTTATTTAGTGTGCT 2953
QY 1921 GTGTCTGCAACCAATGCAAAATCAATTAATTTCTGCTTCTGTGCAATAGAGAGAGTGGG 1980
Db 2954 GTGTCTGCAACCAATGCAAAATCAATTAATTTCTGCTTCTGTGCAATAGAGAGAGTGGG 3013
QY 1981 AGACTTTCTCAGAGTGGGAAGGAGTAATGTAAGGSCATTAATGAAGTCAAAATCTT 2040
Db 3014 AGACTTTCTCAGAGTGGGAAGGAGTAATGTAAGGSCATTAATGAAGTCAAAATCTT 3073
QY 2041 CGAGCGCATCTAGTCAAGTGAAGTGTGCTGTACGAAGGAGAGAAAAAGAACTTATTT 2100
Db 3074 CGAGCGCATCTAGTCAAGTGAAGTGTGCTGTACGAAGGAGAGAAAAAGAACTTATTT 3133
QY 2101 GGAAATGCAAGAGTTTTTTTATGGAACATAAAAAATTAATGCTTCTGTGACAAATTTGCCGAT 2160

Db 3134 GGAAATGCAAGAGTTTTTTTATGAAACATAAAAAATATGCTTCTGTGACAAATTTGCCGAT 3193
QY 2161 CATGCTGAATGTTGATCGGAAGGAAAAATTTGATATCAACAGTGAATAATTAATTTGAATAT 2220
Db 3194 CATGCTGAATGTTGATCGGAAGGAAAAATTTGATATCAACAGTGAATAATTAATTTGAATAT 3253
QY 2221 AAAAAATCCTTCAAAAAATGGCAAGTCTGTTATTTGATTAATTTAGAACTTTTAAAGAGACT 2280
Db 3254 AAAAAATCCTTCAAAAAATGGCAAGTCTGTTATTTGATTAATTTAGAACTTTTAAAGAGACT 3313
QY 2281 TTGGAAAAAGAAAGCAAAATCTCAGAAATATGATCGAAAGATATTGAATCTATTGAAAAA 2340
Db 3314 TTGGAAAAAGAAAGCAAAATCTCAGAAATATGATCGAAAGATATTGAATCTATTGAAAAA 3373
QY 2341 TTATTTGAATGCAATTTTCAGAAAAATTTGATGGAACCCGAGCTTTTACTAATGTGTGA 2400
Db 3374 TTATTTGAATGCAATTTTCAGAAAAATTTGATGGAACCCGAGCTTTTACTAATGTGTGA 3433
QY 2401 AGAATGACAATTAATTTCTCCGATGGAACCTTCAAAAAAGGAATGCTATAGAAATTTGCA 2460
Db 3434 AGAATGACAATTAATTTCTCCGATGGAACCTTCAAAAAAGGAATGCTATAGAAATTTGCA 3493
QY 2461 AACTATGTTCAAGGAGAAATGAAAAATTTAGAGAAAAATTTACCGAAAGATTTTAAAGCT 2520
Db 3494 AACTATGTTCAAGGAGAAATGAAAAATTTAGAGAAAAATTTACCGAAAGATTTTAAAGCT 3553
QY 2521 TTTTCAGAAAGATGATGAGTGAATGATTAAGAAATCTTGAAATTTTACAGAGTAGAAAT 2580
Db 3554 TTTTCAGAAAGATGATGAGTGAATGATTAAGAAATCTTGAAATTTTACAGAGTAGAAAT 3613
QY 2581 TATGCAATTTTTCACACTTTTATCTCTCCGAGCTAATGGAAGAAAGATGTTCTCTCT 2640
Db 3614 TATGCAATTTTTCACACTTTTATCTCTCCGAGCTAATGGAAGAAAGATGTTCTCTCT 3673
QY 2641 GTGGAGAGAGCTGTTCTGTTGGGTAGAACAGGAAATTTATAGCAAGTATCCGTTGAAAA 2700
Db 3674 GTGGAGAGAGCTGTTCTGTTGGGTAGAACAGGAAATTTATAGCAAGTATCCGTTGAAAA 3733
QY 2701 GGAAGCTAACTTGTGCAAAAAAAGATTTAATTAATTAAGCTATCAATTAAGCAGAAACA 2760
Db 3734 GGAAGCTAACTTGTGCAAAAAAAGATTTAATTAATTAAGCTATCAATTAAGCAGAAACA 3793
QY 2761 GTGAATTTTATGTTGGAATATTTGCACTTCGAGAACACAGTACATCCGGAAGTCAAGTCGA 2820
Db 3794 GTGAATTTTATGTTGGAATATTTGCACTTCGAGAACACAGTACATCCGGAAGTCAAGTCGA 3853
QY 2821 GGAAGATTTAATGTTCAAAATCGAAAAATTCAGCTATGCTAGAAAGCTTAAAGAAAAAGCT 2880
Db 3854 GGAAGATTTAATGTTCAAAATCGAAAAATTCAGCTATGCTAGAAAGCTTAAAGAAAAAGCT 3913
QY 2881 GAATTAATCAGAGAAATAATTAATGCAATGCAATGCAAGATTTTTCATGTAGCGGGA 2940
Db 3914 GAATTAATCAGAGAAATAATTAATGCAATGCAATGCAAGATTTTTCATGTAGCGGGA 3973
QY 2941 TCTTTTAAATGTGTGCTCAGGTGGGAATGCAATCAATTTGAAAGCTTAAATTAAGTGAAGCT 3000
Db 3974 TCTTTTAAATGTGTGCTCAGGTGGGAATGCAATCAATTTGAAAGCTTAAATTAAGTGAAGCT 4033
QY 3001 ATCAGTAAGGCAAGAGTTTCAATGATGCAAGCAATTTTGAAGAGCTTAAATTAAGTGAAGCT 3060
Db 4034 ATCAGTAAGGCAAGAGTTTCAATGATGCAAGCAATTTTGAAGAGCTTAAATTAAGTGAAGCT 4093
QY 3061 GCTTTTAAACAGTAAAGATGATATTTCTGTTTGAATGCTGCGGTTCAACCGGAATCGGA 3120
Db 4094 GCTTTTAAACAGTAAAGATGATATTTCTGTTTGAATGCTGCGGTTCAACCGGAATCGGA 4153
QY 3121 ACGAAAAATGCGCGCGGTGCGGTTGCTGTGGCGTAAATGATTAATTTTCAACAA 3180
Db 4154 ACGAAAAATGCGCGCGGTGCGGTTGCTGTGGCGTAAATGATTAATTTTCAACAA 4213
QY 3181 GCTTCAATGAGATTAATGACAAAGCAAAAGTAAATATGATTAAGATTAAGTGAATGA 3240

Db 4214 GCTTCATTGAAGATTAATGACGAAAGACAAAGTAAATATGATTAAGATTAAGATGATGAA 4273
Qy 3241 GTAACAGTAACTGCGGAAATCTTTAGAAAGTATGCAAAAAACGACCGGAAACAATCAACAGT 3300
Db 4274 GTAAACAGTAACTGCGGAAATCTTTAGAAAGTATGCAAAAAACGACCGGAAACAATCAACAGT 4333
Qy 3301 ATTTCTGTTGCCGAGAGAAATTAATAAGTTGGAAGTAAACCGAGTGAAGAAAAACCGAA 3360
Db 4334 ATTTCTGTTGCCGAGAGAAATTAATAAGTTGGAAGTAAACCGAGTGAAGAAAAACCGAA 4393
Qy 3361 TCAGAGAAAGAAACGACGAGAGATTTTGTGCAAAATCGGAAACAAAGTGAAGCTGTATAAA 3420
Db 4394 TCAGAGAAAGAAACGACGAGAGATTTTGTGCAAAATCGGAAACAAAGTGAAGCTGTATAAA 4453
Qy 3421 AATAAATTTACGGATAGTATGATTCATTTACAGAAAAATTTACAAATTTACATTTCTGAA 3480
Db 4454 AATAAATTTACGGATAGTATGATTCATTTACAGAAAAATTTACAAATTTACATTTCTGAA 4513
Qy 3481 GGAGTAAAAAAAGCGGGGAATCTTCCTTCGAAAGTTTCTCATACCTCCGATTAAGAGACG 3540
Db 4514 GGAGTAAAAAAAGCGGGGAATCTTCCTTCGAAAGTTTCTCATACCTCCGATTAAGAGACG 4573
Qy 3541 TCTTTCAGTTTGGAGCTTCTGAAAGTGTTCCTTTCATATATTTTAAAAAGAAACATCT 3600
Db 4574 TCTTTCAGTTTGGAGCTTCTGAAAGTGTTCCTTTCATATATTTTAAAAAGAAACATCT 4633
Qy 3601 GCTGTCTGATGAGATTAAGATTAATTTGAAGGAGCAAAATTAAGAGTGAAGTGAAGT 3660
Db 4634 GCTGTCTGATGAGATTAAGATTAATTTGAAGGAGCAAAATTAAGAGTGAAGTGAAGT 4693
Qy 3661 TCTTCTGATTCATCTTTTGTGAGACATGAGGCGGATCTGCTGCACTTCAGTGAATCAT 3720
Db 4694 TCTTCTGATTCATCTTTTGTGAGACATGAGGCGGATCTGCTGCACTTCAGTGAATCAT 4753
Qy 3721 ATTGGAAGTGAAGAAATAGCAACATCACTGCTGTTTATGCTGAGACCGGCTCTGTAATAT 3780
Db 4754 ATTGGAAGTGAAGAAATAGCAACATCACTGCTGTTTATGCTGAGACCGGCTCTGTAATAT 4813
Qy 3781 ATTTCAAGTGAAGAAACAGTCTTGTGTTAAAAATAGTATATCGAAATGCCAATTAATTT 3840
Db 4814 ATTTCAAGTGAAGAAACAGTCTTGTGTTAAAAATAGTATATCGAAATGCCAATTAATTT 4873
Qy 3841 AAAAGTAAATGCTTTGATGAGAGAACTCAAGTAGCAGCAGAGCAAGTTTGGAGCAGTT 3900
Db 4874 AAAAGTAAATGCTTTGATGAGAGAACTCAAGTAGCAGCAGAGCAAGTTTGGAGCAGTT 4933
Qy 3901 AAAAGTAAATGCTTTGATGAGAGAACTCAAGTAGCAGCAGAGCAAGTTTGGAGCAGTT 3960
Db 4934 AAAAGTAAATGCTTTGATGAGAGAACTCAAGTAGCAGCAGAGCAAGTTTGGAGCAGTT 4993
Qy 3961 GTGAACAATGAAGTTTCTGCAAAATCAGAAAAATTAATACAGTAGCAGAGAAATCTGAAAGC 4020
Db 4994 GTGAACAATGAAGTTTCTGCAAAATCAGAAAAATTAATACAGTAGCAGAGAAATCTGAAAGC 5053
Qy 4021 CAAAAAATGATGTTGATGATGCACTGCTTATCAAGCGGACACCCAAAGTGAACAGAGCTTTA 4080
Db 5054 CAAAAAATGATGTTGATGATGCACTGCTTATCAAGCGGACACCCAAAGTGAACAGAGCTTTA 5113
Qy 4081 AATTTTCAAGCTGGAAGAGTCAATGGAATGTAAGGGGCTTACTGTGACTGTTCGCCAATTA 4140
Db 5114 AATTTTCAAGCTGGAAGAGTCAATGGAATGTAAGGGGCTTACTGTGACTGTTCGCCAATTA 5173
Qy 4141 AACCAACAAGTAAATGCTTCTATAGTGTGGAAGATATATCAAGTAAATGAGAGGAGC 4200
Db 5174 AACCAACAAGTAAATGCTTCTTATAGTGTGGAAGATATATCAAGTAAATGAGAGGAGC 5233
Qy 4201 GCAAAAAGCTCTTTTAGCAACCACTCAAGTACCTGCTGCAAGTGAACGAGGAGGACAAAT 4260
Db 5234 GCAAAAAGCTCTTTTAGCAACCACTCAAGTACCTGCTGCAAGTGAACGAGGAGGACAAAT 5293
Qy 4261 AGTTCTGAGCGGAGTATGAGAAATTAACAAGGGGCTGTTTCTGTCAATTAAGATTGACAT 4320
Db 5294 AGTTCTGAGCGGAGTATGAGAAATTAACAAGGGGCTGTTTCTGTCAATTAAGATTGACAT 5353

Qy 4321 GACGTGGAAGCTAGCGTTGATAAATCTTTCATCCAGAGAGCTAAATGAATCAATGTCATT 4380
Db 5354 GACGTGGAAGCTAGCGTTGATAAATCTTTCATCCAGAGAGCTAAATGAATCAATGTCATT 5413
Qy 4381 GCCAAGATGTCGAAGAGAGTCTGATCTAGCAAAAGAAATATCAAGGCTTATCTAAATGGA 4440
Db 5414 GCCAAGATGTCGAAGAGAGTCTGATCTAGCAAAAGAAATATCAAGGCTTATCTAAATGGA 5473
Qy 4441 AAAGATTAATAATTTTAAAGATCGTGATTTAAATCGATGGAATATGTTAATATACG 4500
Db 5474 AAAGATTAATAATTTTAAAGATCGTGATTTAAATCGATGGAATATGTTAATATACG 5533
Qy 4501 AAGGAAACAATGAAAAAGCAAAAGAAAAAGAGAGCGGCTCATTTGAAATGTCCTTTA 4560
Db 5534 AAGGAAACAATGAAAAAGCAAAAGAAAAAGAGAGCGGCTCATTTGAAATGTCCTTTA 5593
Qy 4561 TCGGTTGCTGGAACGAGTAAATCCGCTGAGAGAGTACTATTTGACATACCTGTTAAA 4620
Db 5594 TCGGTTGCTGGAACGAGTAAATCCGCTGAGAGAGTACTATTTGACATACCTGTTAAA 5653
Qy 4621 AATAAATTTTAAAGAGAAATTAAGTGAAGCAATTAAGAGCGGAGAGATTAATAATCAT 4680
Db 5654 AATAAATTTTAAAGAGAAATTAAGTGAAGCAATTAAGAGCGGAGAGATTAATAATCAT 5713
Qy 4681 GCGAAACATGTAATATGAGAGCAAAATCATCTACTGTTGTTGTAATGCGGCTTCTGGA 4740
Db 5714 GCGAAACATGTAATATGAGAGCAAAATCATCTACTGTTGTTGTAATGCGGCTTCTGGA 5773
Qy 4741 CTGCTATCAGCAAAAGATGCTTTTTCAGAAATGGAGATCGAGCATGGAACACTTATCA 4800
Db 5774 CTGCTATCAGCAAAAGATGCTTTTTCAGAAATGGAGATCGAGCATGGAACACTTATCA 5833
Qy 4801 AATGACACGATTTGCAAGAGTGAATTAAGAGAAATTTCTGCTGAATTCCTTAATGTTGAC 4860
Db 5834 AATGACACGATTTGCAAGAGTGAATTAAGAGAAATTTCTGCTGAATTCCTTAATGTTGAC 5893
Qy 4861 GCAAAATTAATTCATTTCTGAGGAGTAAATGTTCCGGGAAACCATTTGCGGCTTCTCTTCAAG 4920
Db 5894 GCAAAATTAATTCATTTCTGAGGAGTAAATGTTCCGGGAAACCATTTGCGGCTTCTCTTCAAG 5953
Qy 4921 GCGGTGAGAGCTGCTTTTTCGAAATTAATCTTTCATTAATAAAACCTGCTTTGATTAACA 4980
Db 5954 GCGGTGAGAGCTGCTTTTTCGAAATTAATCTTTCATTAATAAAACCTGCTTTGATTAACA 6013
Qy 4981 GGAACGAAGTAAATCCCTTTTATGAGAAAGAAATCAAAAGTCAATGTACAAGCTTTGAT 5040
Db 6014 GGAACGAAGTAAATCCCTTTTATGAGAAAGAAATCAAAAGTCAATGTACAAGCTTTGAT 6073
Qy 5041 GATTCTCATATTTAACAACGTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGGCTGGAATC 5100
Db 6074 GATTCTCATATTTAACAACGTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGGCTGGAATC 6133
Qy 5101 GAGAGAAATGATCTGTCAATCGTGTCTGATGAAACGGAAGCTTTAGTATGATTTCT 5160
Db 6134 GAGAGAAATGATCTGTCAATCGTGTCTGATGAAACGGAAGCTTTAGTATGATTTCT 6193
Qy 5161 GAGTTTGAAGAGTAAATGTTCTTTCAATGTAATGCAAAAGATTCAAAAACAATTAATCA 5220
Db 6194 GAGTTTGAAGAGTAAATGTTCTTTCAATGTAATGCAAAAGATTCAAAAACAATTAATCA 6253
Qy 5221 ATTGCCGGAATATGCAAAATGAGAGAAAGCGGCTGAGTGTGAGCAACAGTTGCTCATGCA 5280
Db 6254 ATTGCCGGAATATGCAAAATGAGAGAAAGCGGCTGAGTGTGAGCAACAGTTGCTCATGCA 6313
Qy 5281 AATAATTGAAAAACAATCAATGATATGATTTGTAATAAAAAACAGTAAATTAACAACGCGAAT 5340
Db 6314 AATAATTGAAAAACAATCAATGATATGATTTGTAATAAAAAACAGTAAATTAACAACGCGAAT 6373
Qy 5341 GATCAAGTAAAAAAAATATCAATGTCCTGCAAAAGTTATATCTATGACCAATATCTATA 5400
Db 6374 GATCAAGTAAAAAAAATATCAATGTCCTGCAAAAGTTATATCTATGACCAATATCTATA 6433

OY	5401	GCAGTCGGAGTTGGAGAGCAAAAGAGCCTCTGTGCAGAGAGCTTCTGCAGTACTACC	5460
Db	6434	GCAGTCGGAGTTGGAGAGCAAAAGAGCCTCTGTGCAGAGAGCTTCTGCAGTACTACC	6493
OY	5461	TTGAAATAGACAGTTCTCTCATGTGTGATCAAACTGATATGTGCAAAAGATTTAGAGAA	5520
Db	6494	TTGAAATAGACAGTTCTCTCATGTGTGATCAAACTGATATGTGCAAAAGATTTAGAGAA	6553
OY	5521	GAATAATATGAAATTAAGAAAAGCAAAATGTTAATGTTCTAGCTGAAAATTAAGAGTCAA	5580
Db	6554	GAATAATATGAAATTAAGAAAAGCAAAATGTTAATGTTCTAGCTGAAAATTAAGAGTCAA	6613
OY	5561	GTGTCACAAATGCGACACATGCTTTCCGAGCAAACTGACACAGTGCAGTAGAGAGCTGCA	5640
Db	6614	GTGTCACAAATGCGACACATGCTTTCCGAGCAAACTGACACAGTGCAGTAGAGAGCTGCA	6673
OY	5611	GTACGAGTTAATAAAATTACACAAATTACTTGACATATTAATAAATTAATTAATCTCAAAAT	5700
Db	6674	GTACGAGTTAATAAAATTACACAAATTACTTGACATATTAATAAATTAATTAATCTCAAAAT	6733
OY	5701	GTACGAAATGCTTTGGTTAAAGCAAAATCTCATTCATCTATTAAACAAATTGGAATTGGA	5760
Db	6734	GTACGAAATGCTTTGGTTAAAGCAAAATCTCATTCATCTATTAAACAAATTGGAATTGGA	6793
OY	5761	GCTGAGTTGAGCTGAGAGAGCTGAGAGTGCAGAGTTCGTAGCAGTGAATAGATTGTA	5820
Db	6794	GCTGAGTTGAGCTGAGAGAGCTGAGAGTGCAGAGTTCGTAGCAGTGAATAGATTGTA	6853
OY	5821	AATATATACGATATGACGAATTTAAATCAGGCCAAATTCCTCGGAGGGAAATATCGGAGTT	5880
Db	6854	AATATATACGATATGACGAATTTAAATCAGGCCAAATTCCTCGGAGGGAAATATCGGAGTT	6913
OY	5881	ATTACAGAGTCTGATGCGGTAAATGCTAAATTATGACAGGAACAGTGTGAGAGTGCCCGT	5940
Db	6914	ATTACAGAGTCTGATGCGGTAAATGCTAAATTATGACAGGAACAGTGTGAGAGTGCCCGT	6973
OY	5941	GCAGCAATAGAGAGCTCAACAGATGTGAATGAATTTACAGGATCTACAAAACATATGTA	6000
Db	6974	GCAGCAATAGAGAGCTCAACAGATGTGAATGAATTTACAGGATCTACAAAACATATGTA	7033
OY	6001	AAAGATTCTACAGTATGCTTAAAGAGAAACAGATGATTTATTACTACTCAAGGGCAA	6060
Db	7034	AAAGATTCTACAGTATGCTTAAAGAGAAACAGATGATTTATTACTACTCAAGGGCAA	7093
OY	6061	GTAGATAAGTGTGATAAAGTATTCAAAATCTTAATATTAACGAAGACTTATCACAA	6120
Db	7094	GTAGATAAGTGTGATAAAGTATTCAAAATCTTAATATTAACGAAGACTTATCACAA	7153
OY	6121	AAAAAATAATTAAGTAAATTAAGAAAGATTTGTACCAATAGTCAAGCTACTCATACTTTA	6180
Db	7154	AAAAAATAATTAAGTAAATTAAGAAAGATTTGTACCAATAGTCAAGCTACTCATACTTTA	7213
OY	6181	AAATCTTTATTGGCAAAATGCGCTGTGTCAAGACAAAGCCGAGTGGCAGAGAACTGTTAAT	6240
Db	7214	AAATCTTTATTGGCAAAATGCGCTGTGTCAAGACAAAGCCGAGTGGCAGAGAACTGTTAAT	7273
OY	6241	ATCAACAAGGTTTATGAGAAACAGAGCTCTGTGAGAAAATTTCTATTTAAATGCAAAA	6300
Db	7274	ATCAACAAGGTTTATGAGAAACAGAGCTCTGTGAGAAAATTTCTATTTAAATGCAAAA	7333
OY	6301	CATTATCTGTAAATTCAGAGATTTACAGAAATTCMAATCGAGTATGATGATTTCTGTGGT	6360
Db	7334	CATTATCTGTGTAAATTCAGAGATTTACAGAAATTCMAATCGAGTATGATGATTTCTGTGGT	7393
OY	6361	GTTTGTGAAATATGTAAGATAGAGACTTCTCTGTATACCAATATTTATTAAGAAATTAAC	6420
Db	7394	GTTTGTGAAATATGTAAGATAGAGACTTCTCTGTATACCAATATTTATTAAGAAATTAAC	7453
OY	6421	AAGACAAGATGAGAAAACTACATATGCTGATGAAGGTTTCCGAGAGAAGAGCTGAATTT	6480
Db	7454	AAGACAAGATGAGAAAACTACATATGCTGATGAAGGTTTCCGAGAGAAGAGCTGAATTT	7513
OY	6481	ACAGCAGATTTCTAAGCAAGAAATTTCTCTTTTGGAGTCCGAGTGCAGACCCGGGGTA	6540

Db	7514	ACAGCAGATTCTAAGCAAGAAATTTCTTCTTTTGAAGTCGAGATCGACAGCCGGGTA	7573
Qy	6541	GGAGCCCGAGTGGCAGGAACCGTTTCCGTAATCAATTTGCAAGAAAGCGAAATGAT	6600
Db	7574	GGAGCCCGAGTGGCAGGAACCGTTTCCGTAATCAATTTGCAAGAAAGCGAAATGAT	7633
Qy	6601	GTGGAAGAGCAAAAGATTTTGGTAAAAAAAAGCTGAGTTACAGAAAACGTTATAGTCT	6660
Db	7634	GTGGAAGAGCAAAAGATTTTGGTAAAAAAAAGCTGAGTTACAGAAAACGTTATAGTCT	7693
Qy	6661	GTTCGAAATGGGAAATGCGCGAGTCGAGTGGCTGCCAAAAGAGCTGGAAATTTGAGCAGCA	6720
Db	7694	GTTCGAAATGGGAAATGCGCGAGTCGAGTGGCTGCCAAAAGAGCTGGAAATTTGAGCAGCA	7753
Qy	6721	GTGCGAGTTACCAAGATGATCAACACGAGACAAAGTGAATAATTTCTAAATATATG	6780
Db	7754	GTGCGAGTTACCAAGATGATCAACACGAGACAAAGTGAATAATTTCTAAATATATG	7813
Qy	6781	ACTCGAAACAAAGTTAGATGTTATGCGAATAATGAGTTAAATATAGGTACTGGAATTCGT	6840
Db	7814	ACTCGAAACAAAGTTAGATGTTATGCGAATAATGAGTTAAATATAGGTACTGGAATTCGT	7873
Qy	6841	TCACCCGAGCTGGAAATTTCTTGACCGCGAGTACTGGAAGTGGTTCTGTCAATATAT	6900
Db	7874	TCACCCGAGCTGGAAATTTCTTGACCGCGAGTACTGGAAGTGGTTCTGTCAATATAT	7933
Qy	6901	GCAATATAGGTAGAAACAGATATGATCATATGTAATTCACTCTTCTACTGATGTAAAT	6966
Db	7934	GCAATATAGGTAGAAACAGATATGATCATATGTAATTCACTCTTCTACTGATGTAAAT	7993
Qy	6961	GTAAAGCTTTAATAAATTTGCAATTTCTTGACAGCCGCTGAGAGAGCCGCGAGTCTT	7020
Db	7994	GTAAAGCTTTAATAAATTTGCAATTTCTTGACAGCCGCTGAGAGAGCCGCGAGTCTT	8053
Qy	7021	GCAGAGTTACCGGAGTGGTTCGTTAAACAATAATAGTTCTGATATGCTCGAGTT	7080
Db	8054	GCAGAGTTACCGGAGTGGTTCGTTAAACAATAATAGTTCTGATATGCTCGAGTT	8113
Qy	7081	CACATATCTCTGATTTGACTTCCGTAACGAAATAAGTAAATGTTAAACGCAAAAGAGAA	7140
Db	8114	CACATATCTCTGATTTGACTTCCGTAACGAAATAAGTAAATGTTAAACGCAAAAGAGAA	8173
Qy	7141	AAAAATATTAAGCAAAACAGACGCAATATGCAAGAAATTCGAGAGCAGCAATTTGAGCCAT	7200
Db	8174	AAAAATATTAAGCAAAACAGACGCAATATGCAAGAAATTCGAGAGCAGCAATTTGAGCCAT	8233
Qy	7201	GTCCTTGGTAATTAATTTTGGAAACAGCTGTAAATAATGAAAAAATTCGAGAGAAAAAGGA	7260
Db	8234	GTCCTTGGTAATTAATTTTGGAAACAGCTGTAAATAATGAAAAAATTCGAGAGAAAAAGGA	8293
Qy	7261	ACAGAAAGTTTAAAAAACTTTAGACGAAGTTAAACAAAGAACAGATTAATAATGAT	7320
Db	8294	ACAGAAAGTTTAAAAAACTTTAGACGAAGTTAAACAAAGAACAGATTAATAATGAT	8353
Qy	7321	GCTACGAAAAAAATCTTACATTCAGACAGTATTTCTACAGAAAGATCTTCTGTAAACCG	7386
Db	8354	GCTACGAAAAAAATCTTACATTCAGACAGTATTTCTACAGAAAGATCTTCTGTAAACCG	8413
Qy	7381	GATAGAGAGATATCTCAGAGGAAAGAAATTAACCATTTGGAAGACTTCTGATATATAT	7440
Db	8414	GATAGAGAGATATCTCAGAGGAAAGAAATTAACCATTTGGAAGACTTCTGATATATAT	8473
Qy	7441	GGAAAAAATGTAGATATTAACAACAGAGCAAGAATAATATCACTTCTACTGGTGTG	7500
Db	8474	GGAAAAAATGTAGATATTAACAACAGAGCAAGAATAATATCACTTCTACTGGTGTG	8533
Qy	7501	GGAACTCAGAGTCTTGTCTTCCGATCAGAAAACGTGGCAGTTACAAATATTTAAAGAAAT	7566
Db	8534	GGAACTCAGAGTCTTGTCTTCCGATCAGAAAACGTGGCAGTTACAAATATTTAAAGAAAT	8593
Qy	7561	TCCGAGATTACGTGTGAAAAATTTCTTTTGAAGACGTGAAAAAGTAAATGTTAGATCG	7620

Db 8594 TCCGAGTTACTGTGAAAAATCTTTGTGAAAGCACTGAAAAAGTAAATTTAGATCG 8653
Qy 7621 GATATTACAGAAATGTTGCTTTAA CAGCATATCAAGTCTCTAGAGCATTTGGAAATA 7680
Db 8654 GATATTACAGAAATGTTGCTTTAA CAGCATATCAAGGTCCTGTAGAGCATTTGGAAATA 8713
Qy 7681 GAGAGTGCCTATGCAAGAAATTAATTTCTAATGGAAGATCAAAATATCAATTTAAAAATCT 7740
Db 8714 GAGAGTGCCTATGCAAGAAATTAATTTCTAATGGAAGATCAAAATATCAATTTAAAAATCT 8773
Qy 7741 AAGCTATTAGAAAAAATATGATGTTATTTGTAAGAAATTAATTCGAAATTTGAGCGGAA 7800
Db 8774 AAGCTATTAGAAAAAATATGATGTTATTTGTAAGAAATTAATTCGAAATTTGAGCGGAA 8833
Qy 7801 GCAAAAAGATTAACCGTAGAGCGGTAGCTGCGGAGCCATTATCTCAAAAGCAAGAAT 7860
Db 8834 GCAAAAAGATTAACCGTAGAGCGGTAGCTGCGGAGCCATTATCTCAAAAGCAAGAAT 8893
Qy 7861 GAAATGAAATTCAGAGTTGAAATTTGAGAAAGATATTTTCATGAAAGAAATAGAGTAAT 7920
Db 8894 GAAATGAAATTCAGAGTTGAAATTTGAGAAAGATATTTTCATGAAAGAAATAGAGTAAT 8953
Qy 7921 AGCCCTTCTAAAGAAATTGAAAGAGAAATCAATGTCAAAAGTGAAAAAGAAACAGAGTG 7980
Db 8954 AGCCCTTCTAAAGAAATTGAAAGAGAAATCAATGTCAAAAGTGAAAAAGAAACAGAGTG 9013
Qy 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGAGGGGCAAGATTAATTTCCGA 8040
Db 9014 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGAGGGGCAAGATTAATTTCCGA 9073
Qy 8041 GCAAAAAGATGCGGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100
Db 9074 GCAAAAAGATGCGGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 9133
Qy 8101 CATGCAAGATTAATGTGAATATGAGAGCAACATTAATTAAGATTAACAGCAGTTCTAA 8160
Db 9134 CATGCAAGATTAATGTGAATATGAGAGCAACATTAATTAAGATTAACAGCAGTTCTAA 9193
Qy 8161 GCAGTAAACAGGTTCTGTATTTGGAGAGAGTTGAGTCAACCAAGCAGAAAGCTACTGCA 8220
Db 9194 GCAGTAAACAGGTTCTGTATTTGGAGAGAGTTGAGTCAACCAAGCAGAAAGCTACTGCA 9253
Qy 8221 GGTAAACATATGTTAGAAAGTTGAGAGAGAAATTTGTTCAACAATTCGATTTGAATGCA 8280
Db 9254 GGTAAACATATGTTAGAAAGTTGAGAGAGAAATTTGTTCAACAATTCGATTTGAATGCA 9313
Qy 8281 ATTTCTAAAGTAGAAGGTTTGATGAAGATAAGTAACTGTAATCTTCTGTAGTATCA 8340
Db 9314 ATTTCTAAAGTAGAAGGTTTGATGAAGATAAGTAACTGTAATCTTCTGTAGTATCA 9373
Qy 8341 GGAAATGAGAGAGATTGCGCGAGCAGAGTGAATACTTCTTACAGCACAAAAGTAACT 8400
Db 9374 GGAAATGAGAGAGATTGCGCGAGCAGAGTGAATACTTCTTACAGCACAAAAGTAACT 9433
Qy 8401 GAAATCCGTAAGTTCGTTTACGAAAGCAAGATTATGAAATATATGATTACCAAAAAATAT 8460
Db 9434 GAAATCCGTAAGTTCGTTTACGAAAGCAAGATTATGAAATATATGATTACCAAAAAATAT 9493
Qy 8461 ATTTGAGAAGTCAATGCTCTTGTCTTAAATGATACAAAGATGAAGCGAATTAAGTACT 8520
Db 9494 ATTTGAGAAGTCAATGCTCTTGTCTTAAATGATACAAAGATGAAGCGAATTAAGTACT 9553
Qy 8521 TTAGCGGTAGCCGAGTGTGATGCACAGAGAGCAAAACAAAGATTTTACAGATCAAAACAG 8580
Db 9554 TTAGCGGTAGCCGAGTGTGATGCACAGAGAGCAAAACAAAGATTTTACAGATCAAAACAG 9613
Qy 8581 TTAACTTCTACACTGTTAAATGAGAGAAAGTATCTCAACTTGTGCAAAAGCTTTGCT 8640
Db 9614 TTAACTTCTACAACTGTTAAATGAGAGAAAGTATCTCAACTTGTGCAAAAGCTTTGCT 9673
Qy 8641 AAAAAATGAATAATTAAGAGAAATGTAAGAGAACTGAGAGAGCTTGTGAGACCGGAAACA 8700
Db 9674 AAAAAATGAATAATTAAGAGAAATGTAAGAGAACTGAGAGAGCTTGTGAGACCGGAAACA 9733

Qy 8701 GCAGCGTTGAAAAATTATACAAAGATCTACAGAGCATTTGTTGACAGAAATTTGGAA 8760
Db 9734 GCAGCGTTGAAAAATTATACAAAGATCTACAGAGCATTTGTTGACAGAAATTTGGAA 9793
Qy 8761 ATTGAGATTAATTTGAAAGCAATTTGCAAGAGATTAATACGATTTGTAAGTCAACGGAAC 8820
Db 9794 ATTGAGATTAATTTGAAAGCAATTTGCAAGAGATTAATACGATTTGTAAGTCAACGGAAC 9853
Qy 8821 GGAACCAAGAGAGTCTTGTGCGAAGAGATGTAATTTCTGTGAAAAATATCAATTTACAGG 8880
Db 9854 GGAACCAAGAGAGTCTTGTGCGAAGAGATGTAATTTCTGTGAAAAATATCAATTTACAGG 9913
Qy 8881 GAAACAAATATCATTCATTTGAAGATTAAGCCAGAAATTTGGAACCGGAAGTAAATGTA 8940
Db 9914 GAAACAAATATCATTCATTTGAAGATTAAGCCAGAAATTTGGAACCGGAAGTAAATGTA 9973
Qy 8941 GATGCTTTGAATGAACCTTGAATGTGATCTTACAGAGAAAAATGCTGCTATGTGGAATT 9000
Db 9974 GATGCTTTGAATGAACCTTGAATGTGATCTTACAGAGAAAAATGCTGCTATGTGGAATT 10033
Qy 9001 GGTATTGGAATTTGATGTAAATTAATGATTAAGAAAAATGTAGAAAGCCAAATCCGA 9060
Db 10034 GGTATTGGAATTTGATGTAAATTAATGATTAAGAAAAATGTAGAAAGCCAAATCCGA 10093
Qy 9061 AGACATGCTATTGTAGAAATCTAGAAAAACAAGATATCAAGCATTTTACAAAGCAAAA 9120
Db 10094 AGACATGCTATTGTAGAAATCTAGAAAAACAAGATATCAAGCATTTTACAAAGCAAAA 10153
Qy 9121 GTAAATATTTCTTGGAAAAAGAGAGAGCTGCACTGCACTGCAATATGCAATGTACACTT 9180
Db 10154 GTAAATATTTCTTGGAAAAAGAGAGAGCTGCACTGCACTGCAATATGCAATGTACACTT 10213
Qy 9181 TCCATGAGATGATTAATTAATTTTGGCAAGCAGTATGATCTTCTCAATTAATAAC 9240
Db 10214 TCCATGAGATGATTAATTAATTTTGGCAAGCAGTATGATCTTCTCAATTAATAAC 10273
Qy 9241 AAAAAATCAAAAAATTAATTAATTTAGTATGATCAAGTGAATGCAATGTAATGTTGAT 9300
Db 10274 AAAAAATCAAAAAATTAATTAATTTAGTATGATCAAGTGAATGCAATGTAATGTTGAT 10333
Qy 9301 GGGGTGCTGAAGCAAGAGGTGCGAGAGCCAAACGCAAGTATGTTAAAGATCAAAATA 9360
Db 10334 GGGGTGCTGAAGCAAGAGGTGCGAGAGCCAAACGCAAGTATGTTAAAGATCAAAATA 10393
Qy 9361 AATGAACATAATTAATTTGATTTAGCAGAGAAATTAATAAGAGGAAACATCAATGTA 9420
Db 10394 AATGAACATAATTAATTTGATTTAGCAGAGAAATTAATAAGAGGAAACATCAATGTA 10453
Qy 9421 TATGCCGATATGATTAATAATTTATATTAAGTAAAGCAAAATTTCTAAGCTATTGCCGAT 9480
Db 10454 TATGCCGATATGATTAATAATTTATATTAAGTAAAGCAAAATTTCTAAGCTATTGCCGAT 10513
Qy 9481 GCCAAAGTCAATGCTGCAAGCTGCTTCCGCAACTGCACTATTGGAAGAAAAATGAAGTAAA 9540
Db 10514 GCCAAAGTCAATGCTGCAAGCTGCTTCCGCAACTGCACTATTGGAAGAAAAATGAAGTAAA 10573
Qy 9541 TTTAATTAATGCGATCCGGAATTTAAAAATATATGCAAGATTTGAAAGGAAAGCTAAT 9600
Db 10574 TTTAATTAATGCGATCCGGAATTTAAAAATATATGCAAGATTTGAAAGGAAAGCTAAT 10633
Qy 9601 AAAAAAGCTGCGATGATCTTAATCAAGTATGATGTAACGATTAATATATATATGCGAT 9660
Db 10634 AAAAAAGCTGCGATGATCTTAATCAAGTATGATGTAACGATTAATATATATATGCGAT 10693
Qy 9661 TCTTCTGAAAAAGCATACAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGAGAA 9720
Db 10694 TCTTCTGAAAAAGCATACAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGAGAA 10753
Qy 9721 AAATGA 9726
Db 10754 AAATGA 10759

RESULT 3
US-09-841-786-10
Sequence 10, Application US/09941786
Patent No. 6669940
GENERAL INFORMATION:
APPLICANT: NAGARAJA, T. G.
APPLICANT: STEWART, GEORGE C.
APPLICANT: NARAYANAN, SANDEEV K.
APPLICANT: CHENGAPPA, M. M.
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
FILE REFERENCE: VACCINE AND PREPARATION THEREOF
CURRENT APPLICATION NUMBER: US/09/841,786
CURRENT FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 09/558,257
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 2780
TYPE: DNA
ORGANISM: Fusobacterium necrophorum
US-09-841-786-10

Query Match 28.6%; Score 2780; DB 3; Length 2780;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 2780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 GGGAGATTATTAAGAAAGGAGGAGAAAGAACTTATACCTCTTAAAGTTATCA 978
DB 1 GGGAGATTATTAAGAAAGGAGGAGAAAGAAAGAACTTATACCTCTTAAAGTTATCA 60
QY 979 GATGGAAGCTTCCGTAAAGATTAATAAGAAAGTCAAGAAAGATGTTGACAT 1038
DB 61 GATGGAAGCTTCCGTAAAGATTAATAAGAAAGTCAAGAAAGATGTTGACAT 120
QY 1039 ACAGCTGAAGCAAAAGATTTCTATGATGCACTTTAGTTACTAGCTTGCAAGCACTCT 1098
DB 121 ACAGCTGAAGCAAAAGATTTCTATGATGCACTTTAGTTACTAGCTTGCAAGCACTCT 180
QY 1099 TTTAGCTTTGTTACAGGTTCTATTTCTCCATCAATTTAATGATTTTAAAGTTATG 1158
DB 181 TTTAGCTTTGTTACAGGTTCTATTTCTCCATCAATTTAATGATTTTAAAGTTATG 240
QY 1159 ACAGTAAGTCCAGTGTCTTATTTGAAAGATGCAAACTGCAAGCAAGAAAGAAAG 1218
DB 241 ACAGTAAGTCCAGTGTCTTATTTGAAAGATGCAAACTGCAAGCAAGAAAGAAAG 300
QY 1219 GCAAAATATTCATTTCTTACGTGAGTAAGCACTATGAGAGCACTTCTCATTA 1278
DB 301 GCAAAATATTCATTTCTTACGTGAGTAAGCACTATGAGAGCACTTCTCATTA 360
QY 1279 AAAAATACCAATTTATTTTGAAGAAAGCAATGGAAGCTTCTAGTATCGAAGGGGA 1338
DB 361 AAAAATACCAATTTATTTTGAAGAAAGCAATGGAAGCTTCTAGTATCGAAGGGGA 420
QY 1339 TATATTTGCAAAAGATTAATCCAAATGTAATGTAAGAGAAAGTAAATGCAAGGA 1398
DB 421 TATATTTGCAAAAGATTAATCCAAATGTAATGTAAGAGAAAGTAAATGCAAGGA 480
QY 1399 AGAGCAGATTTAATCTTCAAAATCTGAAATATGATGATGCTTCTGTTCTGTTGAACG 1458
DB 481 AGAGCAGATTTAATCTTCAAAATCTGAAATATGATGATGCTTCTGTTCTGTTGAACG 540
QY 1459 ATAGAGATTTCAATTAAGAGCTTTCTGATGTTGTTGAGAGAGAAAGTAAATCT 1518
DB 541 ATAGAGATTTCAATTAAGAGCTTTCTGATGTTGTTGAGAGAGAAAGTAAATCT 600
QY 1519 TCCGTCAAGATGCTTAAGAGCAAAAGTAAATGCAAGAGATGATGTAATGTCAGA 1578
DB 601 TCCGTCAAGATGCTTAAGAGCAAAAGTAAATGCAAGAGATGATGTAATGTCAGA 660

QY 1579 AGTGAAGGATTAATTCATTCAGAGCTGCTGTAAGAAAGTGAATGGGAGATAGTGAAT 1638
DB 661 AGTGAAGGATTAATTCATTCAGAGCTGCTGTAAGAAAGTGAATGGGAGATAGTGAAT 720
QY 1639 GGGGTTGTGGCTGCAAAATATTTCTAACTATAATGCTTCTCCGTAATAGATGATGGA 1698
DB 721 GGGGTTGTGGCTGCAAAATATTTCTAACTATAATGCTTCTCCGTAATAGATGATGGA 780
QY 1699 TATCTACATGCCAAGACGACTAAATGAGAGGCTCATACATTAATTAATGCTGT 1758
DB 781 TATCTACATGCCAAGACGACTAAATGAGAGGCTCATACATTAATTAATGCTGT 840
QY 1759 CTGCAAGCAGATCGATTTGGGAAGCTTCAAGTTATGAATGATCAGTTATGAATCA 1818
DB 841 CTGCAAGCAGATCGATTTGGGAAGCTTCAAGTTATGAATGATCAGTTATGAATCA 900
QY 1819 GGTCAATCAATCAATTTTATGATGCAATAAACAGCGTTTGAAGAGACAGTGTCAAT 1878
DB 901 GGTCAATCAATCAATTTTATGATGCAATAAACAGCGTTTGAAGAGACAGTGTCAAT 960
QY 1879 GAGGAAATTAAGAAATTAAGCTAACGAACTTATTAATGATGCTGCTGCAACCAATAGA 1938
DB 961 GAGGAAATTAAGAAATTAAGCTAACGAACTTATTAATGATGCTGCTGCAACCAATAGA 1020
QY 1939 AATCAATTAATTTCTGCTTGTGGCAATAGAGAGAGTGAAGACTTCTTCAAGAGTG 1998
DB 1021 AATCAATTAATTTCTGCTTGTGGCAATAGAGAGAGTGAAGACTTCTTCAAGAGTG 1080
QY 1999 GAAGGAGATTAATGTAAGGCAATTAATGAGCTCAAAATCTTGAAGCACTACGTCAGT 2058
DB 1081 GAAGGAGATTAATGTAAGGCAATTAATGAGCTCAAAATCTTGAAGCACTACGTCAGT 1140
QY 2059 GGAAGTGTGCTGTACGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2118
DB 1141 GGAAGTGTGCTGTACGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 2119 TATGAAATCTAATAAATAATGCTTCTGTAACAATGCGCATAGCTGAATTTGGATCG 2178
DB 1201 TATGAAATCTAATAAATAATGCTTCTGTAACAATGCGCATAGCTGAATTTGGATCG 1260
QY 2179 GAAGGAAATTTGATATCAACAGTGAATAAATAATGGAATATTAATTAATCTTCAAAATG 2238
DB 1261 GAAGGAAATTTGATATCAACAGTGAATAAATAATGGAATATTAATTAATCTTCAAAATG 1320
QY 2239 GCAAAAGTCTGTATGATTAATTAAGACTTTTAAAGAGAGCTTTTGAAGAAAGAGAA 2298
DB 1321 GCAAAAGTCTGTATGATTAATTAAGACTTTTAAAGAGAGCTTTTGAAGAAAGAGAA 1380
QY 2299 ACTCGAATATGATCCGAAGATTAATGAACTATTTGAAGAAATTAATGAAATGCAATTTCA 2358
DB 1381 ACTCGAATATGATCCGAAGATTAATGAACTATTTGAAGAAATTAATGAAATGCAATTTCA 1440
QY 2359 GAAAAATGATGAG 2418
DB 1441 GAAAAATGATGAG 1500
QY 2419 CCGGATGGAATCTTCAAAACAGGAACTGCTATAGAAATGGAATGCAATGTTGAGAGAA 2478
DB 1501 CCGGATGGAATCTTCAAAACAGGAACTGCTATAGAAATGGAATGCAATGTTGAGAGAA 1560
QY 2479 ATGAAATTAATGAG 2538
DB 1561 ATGAAATTAATGAG 1620
QY 2539 GGAAGTGAATTAAGAAATTTGAATTTTACAGAGTGAAGAAATTAATGCAATTTTCAACT 2598
DB 1621 GGAAGTGAATTAAGAAATTTGAATTTTACAGAGTGAAGAAATTAATGCAATTTTCAACT 1680
QY 2599 TTTACTCTTCCGAGCTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2658
DB 1681 TTTACTCTTCCGAGCTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 2659 TGGGTAAGACAGAGAAATTAATGCAAGATTCCTTGAAGAGAGAGAGAGAGAGAGAG 2718

```
Db 1741 TGGGTGAGACAGAGAAATTAATAGCAAGGATCCGTTGGAAAAAGAGCTAAATCTGCTGCA 1800
Qy 2719 AAAAAAGATTTAAATATAAAGCTATCAATAAAGCAAGGTAATTTAGTTGGAAT 2778
Db 1801 AAAAAAGATTTAAATATAAAGCTATCAATAAAGCAAGGTAATTTAGTTGGAAT 1860
Qy 2779 ATTGACCTTGCAGAAAGCAGTACATCCGAGAGTGCAGTGCAGAAAGATTTAAATGTTCA 2838
Db 1861 ATTGACCTTGCAGAAAGCAGTACATCCGAGAGTGCAGTGCAGAAAGATTTAAATGTTCA 1920
Qy 2839 AGATGAAAAAATTCAGCTATCCGTAGAAGCTAAAGAAAAAGCTGAATTAACAGAGAAAAAT 2898
Db 1921 AGATGAAAAAATTCAGCTATCCGTAGAAGCTAAAGAAAAAGCTGAATTAACAGAGAAAAAT 1980
Qy 2899 ATTAATGACAGATGCAATGGAACAGACTTTTCAATGAGCGGGATCTTTTAATGTTGCTCA 2958
Db 1981 ATTAATGACAGATGCAATGGAACAGACTTTTCAATGAGCGGGATCTTTTAATGTTGCTCA 2040
Qy 2959 GGTGGGAAATGCAATCAATGGAATGGGAAGTTATAGTGAAGTATCAGTAAGGCAAGATT 3018
Db 2041 GGTGGGAAATGCAATCAATGGAATGGGAAGTTATAGTGAAGTATCAGTAAGGCAAGATT 2100
Qy 3019 TCCATTGATGACGAAGCATATTTGAAAGCTAATAAAAAATGCTTTAAACAGTAAGAT 3078
Db 2101 TCCATTGATGACGAAGCATATTTGAAAGCTAATAAAAAATGCTTTAAACAGTAAGAT 2160
Qy 3079 GATTACTTCGTTGGAAATGCTGCGGTTGAGCGGGAAATCGGAAACGAAAAATCGCGCGGTC 3138
Db 2161 GATTACTTCGTTGGAAATGCTGCGGTTGAGCGGGAAATCGGAAACGAAAAATCGCGCGGTC 2220
Qy 3139 GGGGTTGCTGTTGCGGTAATGATTAATGATTAATTTCAAAACAAGCTTCCATTGAAGATAT 3198
Db 2221 GGGGTTGCTGTTGCGGTAATGATTAATGATTAATTTCAAAACAAGCTTCCATTGAAGATAT 2280
Qy 3199 GACGAAAGCAAGTAATATAATGATTAAGATTAAGATGATGAAAGTAACTGCGGAA 3258
Db 2281 GACGAAAGCAAGTAATATAATGATTAAGATTAAGATGATGAAAGTAACTGCGGAA 2340
Qy 3259 TCTTTAGAAGTATGATGCAAAAAAGCAACGAAACAATCAAGTATTTCTGTTCCGGAAGA 3318
Db 2341 TCTTTAGAAGTATGATGCAAAAAAGCAACGAAACAATCAAGTATTTCTGTTCCGGAAGA 2400
Qy 3319 ATTAATAAGTATGAAAGTAAACGAGTGAAGAAAAACGAAATCAGAAAGAAAGACAGAG 3378
Db 2401 ATTAATAAGTATGAAAGTAAACGAGTGAAGAAAAACGAAATCAGAAAGAAAGACAGAG 2460
Qy 3379 GGAATTTTGGCAAAATCGGAAACAAAGTGAAGCTCTGTAATAATTAATTAATTAATTAAT 3438
Db 2461 GGAATTTTGGCAAAATCGGAAACAAAGTGAAGCTCTGTAATAATTAATTAATTAATTAAT 2520
Qy 3439 ATGATTCATTAACGAAAAAATTAACAATTACATTTTCGAGAGGATAAAAAACCGGGG 3498
Db 2521 ATGATTCATTAACGAAAAAATTAACAATTACATTTTCGAGAGGATAAAAAACCGGGG 2580
Qy 3499 AATCTTCTCCGAAAGTTCATCTCCGATTAAGGACGCTCTTTCAGTTGGAGACT 3558
Db 2581 AATCTTCTCCGAAAGTTCATCTCCGATTAAGGACGCTCTTTCAGTTGGAGACT 2640
Qy 3559 TCTGGAAGTGTCTTTCAATATAATTAATAAAGAAACATCTGCTGCTAGATGAGATA 3618
Db 2641 TCTGGAAGTGTCTTTCAATATAATTAATAAAGAAACATCTGCTGCTAGATGAGATA 2700
Qy 3619 AAGATAAATTTGAAGGAGCAATAAAAAGTGAAGGATCTTCTTCGATTTCTTCTTTT 3678
Db 2701 AAGATAAATTTGAAGGAGCAATAAAAAGTGAAGGATCTTCTTCGATTTCTTCTTTT 2760
Qy 3679 GTTGAAGCATGGGCGGATC 3698
Db 2761 GTTGAAGCATGGGCGGATC 2780
```

```
US-09-841-786-13
; Sequence 13, Application US/09841786
; Patent No. 6669940
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2322
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-13
```

```
Query Match 23.9%; Score 2322; DB 3; Length 2322;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 7405 GGAATTAAGCATTTGGAACACCTTCTGATATTAATGGAATAATGTTAGATTTTACACA 7464
Db 1 GGAATTAAGCATTTGGAACACCTTCTGATATTAATGGAATAATGTTAGATTTTACACA 60
Qy 7465 GAGACAAGATTAATATCACTTCTACCTGCTGTTGGAACTGCAAGCTCTTGCCGA 7524
Db 61 GAGACAAGATTAATATCACTTCTACCTGCTGTTGGAACTGCAAGCTCTTGCCGA 120
Qy 7525 TCAGAACAGTGGCAGTTTACAAATATTAAAGAAATTCGAGATTACTGTTGAAATTTCT 7584
Db 121 TCAGAACAGTGGCAGTTTACAAATATTAAAGAAATTCGAGATTACTGTTGAAATTTCT 180
Qy 7585 TTTGTGAAGAGCTGGAATAATGTTAGATTCGATATTAACAGAAATGTTGCTTTA 7644
Db 181 TTTGTGAAGAGCTGGAATAATGTTAGATTCGATATTAACAGAAATGTTGCTTTA 240
Qy 7645 ACAGCATATCAAGGCTCTGTAGAGCATTTGGAATTAGAGCTGCTATGCAGAATTTAAT 7704
Db 241 ACAGCATATCAAGGCTCTGTAGAGCATTTGGAATTAGAGCTGCTATGCAGAATTTAAT 300
Qy 7705 TCTAATGGAAGATCAAAATATCAATTAATAATTTCTAAGCTATTAGAAAAAATATTGAT 7764
Db 301 TCTAATGGAAGATCAAAATATCAATTAATAATTTCTAAGCTATTAGAAAAAATATTGAT 360
Qy 7765 GTTATTTGTAAGATTAATTCGAAATTTGAGACCGGAAGCAAAAGATTAAACCGTAGAGCG 7824
Db 361 GTTATTTGTAAGATTAATTCGAAATTTGAGACCGGAAGCAAAAGATTAAACCGTAGAGCG 420
Qy 7825 GTAGCTGCCGAGGCATTTATCTCAAAAGCAAAAGATGAATGAATTCAGAGGTGAATTT 7884
Db 421 GTAGCTGCCGAGGCATTTATCTCAAAAGCAAAAGATGAATGAATTCAGAGGTGAATTT 480
Qy 7885 GAGAAAGATTTTCAATGAAGAAATAGAGTAATCTAGCCCTTCTAAAGAAATTTGAAAGA 7944
Db 481 GAGAAAGATTTTCAATGAAGAAATAGAGTAATCTAGCCCTTCTAAAGAAATTTGAAAGA 540
Qy 7945 GAAATCAATGTCAGAGTGAAGAAAGAAAGCAAGATGACTGTGAATTCGAGAGCTTCT 8004
Db 541 GAAATCAATGTCAGAGTGAAGAAAGAAAGCAAGATGACTGTGAATTCGAGAGCTTCT 600
Qy 8005 GTAGAGCAGTGAAGCAGAGGCGAGAAATTTCTCGAAGCAAAAGATGCGGAAAGCTCTTAT 8064
Db 601 GTAGAGCAGTGAAGCAGAGGCGAGAAATTTCTCGAAGCAAAAGATGCGGAAAGCTCTTAT 660
Qy 8065 TTGAAAGTTAGTCAAAATCCGAAAGAGTATTTTTCATGCGAATTAATGTAATGAGAA 8124
Db 8124
```

Db	661	TTGAAAGTGTACTCAAATATCCGGAAGAAGTATTTTTCATGCAGATATATGTGAATATGCA	720
Qy	8125	GCAACACATAAATGAAGTATACAGCAGTTTCTTAAAGCATACAGGTTCTGTATTGGGA	8184
Db	721	GCAACACATAAATGAAGTATACAGCAGTTTCTTAAACAGTATACAGGTTCTGTATTGGGA	780
Qy	8185	GGAGTTGCAATCACCAAGGCAAGAGCTATGCTGTCAGGTAAACCTATGTGTGAAGTTGAG	8244
Db	781	GGAGTTGGAAGTCAACCAAGGCAAGCTATGCTGTCAGGTAAACCTATGTGTGAAGTTGAG	840
Qy	8245	GAGGAATTTTGTTCAGAACCAATTCGATTTGAATGCAATTTCTTAAAGTGAAGGTTTGGAT	8304
Db	841	GAGGAATTTTGTTCAGAACCAATTCGATTTGAATGCAATTTCTTAAAGTGAAGGTTTGGAT	900
Qy	8305	GAAGTAAAGTAACTGCTTAAATCTTCTGTAGTATCAGGAAATGAGAGGAAATTTGCCGA	8364
Db	901	GAAGTAAAGTAACTGCTTAAATCTTCTGTAGTATCAGGAAATGAGAGGAAATTTGCCGA	960
Qy	8365	GCAGGAGTGAATTACTTCTTACAGCACCAAGTATATCTGAATCCGTAATTCGTTTACGAAG	8424
Db	961	GCAGGAGTGAATTACTTCTTACAGCACCAAGTATATCTGAATCCGTAATTCGTTTACGAAG	1020
Qy	8425	CAAGTTATGAAATTAAGATTATACAAAAAATATATTTTCAGAAAGTCAATGCTCTTGCT	8484
Db	1021	CAAGTTATGAAATTAAGATTATACAAAAAATATATTTTCAGAAAGTCAATGCTCTTGCT	1080
Qy	8485	TTAAATGATACAAAGATGAAGCGAATATAGATCTTTAGCGGTAGCGGTGTGCATGCA	8544
Db	1081	TTAAATGATACAAAGATGAAGCGAATATAGATCTTTAGCGGTAGCGGTGTGCATGCA	1140
Qy	8545	CAAGGACAAACAAAGCATTTACGAGATCAAAACAATTAACTTTCACACTGTAAATGGA	8604
Db	1141	CAAGGACAAACAAAGCATTTACGAGATCAAAACAATTAACTTTCACACTGTAAATGGA	1200
Qy	8605	GGAACGATCTCAACTTCGTCGCAAAACCTTGGCTAAATGTAAATTTATGAAATGTA	8664
Db	1201	GGAACGATCTCAACTTCGTCGCAAAACCTTGGCTAAATGTAAATTTATGAAATGTA	1260
Qy	8665	AAAGGAATGGAAGAGCCTTAGTTCGAGACCGGAAACAGCAGCCGTGTGAATTTATACAAG	8724
Db	1261	AAAGGAATGGAAGAGCCTTAGTTCGAGACCGGAAACAGCAGCCGTGTGAATTTATACAAG	1320
Qy	8725	AGTACTACAGAGCATTTGCTTGCAGGAAATTTGGAAATTTGAGATTAATTTGAAAACGATT	8784
Db	1321	AGTACTACAGAGCATTTGCTTGCAGGAAATTTGGGAAATTTGAGATTAATTTGAAAACGATT	1380
Qy	8785	GCAAGATTAATACATTTGTGAAGATCAACGGAGACGGAACCAAGAGCTTGTGTGGA	8844
Db	1381	GCAAGATTAATACATTTGTGAAGATCAACGGAGACGGAACCAAGAGCTTGTGTGGA	1440
Qy	8845	AAGATGTGATTTTCGTGTGAAAAATACATTTTCAAGGGGAAACAAATATCATTCATTGGAAGAT	8904
Db	1441	AAGATGTGATTTTCGTGTGAAAAATACATTTTCAAGGGGAAACAAATATCATTCATTGGAAGAT	1500
Qy	8905	AAAGCCAGAAATTTGTGGAACCGGAAGTGTAAATGTAGATGCTTTGATGAATCTTGATGTA	8964
Db	1501	AAAGCCAGAAATTTGTGGAACCGGAAGTGTAAATGTAGATGCTTTGATGAATCTTGATGTA	1560
Qy	8965	GATCTTACAGAAAAAGTGTGCTATGTGTGGAATTTGGTATTTGGAAATGTTGATGTAAAT	9024
Db	1561	GATCTTACAGAAAAAGTGTGCTATGTGTGGAATTTGGTATTTGGAAATGTTGATGTAAAT	1620
Qy	9025	AATGTGATTTAAAGAAAAATGTAGAACCAAAATCGGAAGACATGTATTTGTGAAATCTACT	9084
Db	1621	AATGTGATTTAAAGAAAAATGTAGAACCAAAATCGGAAGACATGTATTTGTGAAATCTACT	1680
Qy	9085	GGAAAAACAAGATATCAAGCATTTTACAAAGACCAAAAGTAAATATTCTTGGAAAAAGAGAC	9144
Db	1681	GGAAAAACAAGATATCAAGCATTTTACAAAGACCAAAAGTAAATATTCTTGGAAAAAGAGAC	1740
Qy	9145	GCTGCAGCTGCAGCTGCATATTCGAAATGTACATTTTCCATGTAGATGGAATTTTAAAT	9204
Db	1741	GCTGCAGCTGCAGCTGCATATTCGATGTACATTTTCCAAATGAATGGAATTTTAAAT	1800

OY		9205	TTGGCAAGACGATGTGCATCTTCTGCATTATAACCCAAAAATTCAAAATTAATTAATACT	9264
Dd		1801	TTGGCAAGACGATGTGCATCTTCTGCATTATAACCCAAAAATTCAAAATTAATTAATACT	1860
OY		9265	TTAGCATCAGTAGTGGAATCGAATGTGAATGTTTCATGGGTGGCTGAAGCAAGAGTGCA	9324
Dd		1861	TTAGCATCAGTAGTGGAATCGAATGTGAATGTTTCATGGGTGGCTGAAGCAAGAGTGCA	1920
OY		9325	GGACCCAAAGCGACAGTTAGTGTAAAGAATCAAAATAATVGAACATAATTAATGTTGATTTA	9384
Dd		1921	GGACCCAAAGCGACAGTTAGTGTAAAGAATCAAAATAATVGAACATAATTAATGTTGATTTA	1980
OY		9385	GCAGCAAAAAATTTAAACAAGAGGGAACAATCAATCTATATGCCGGAATGTGAATAAAATTAAT	9444
Dd		1981	GCAGCAAAAAATTTAAACAAGAGGGAACAATCAATCTATATGCCGGAATGTGAATAAAATTAAT	2040
OY		9445	AATATAAGTAAGACAAATTTCTAAGGCTATTGCCGATGCCAAAAGTCATGCTGCAGCTCT	9504
Dd		2041	AATATAAGTAAGACAAATTTCTAAGGCTATTGCCGATGCCAAAAGTCATGCTGCAGCTCT	2100
OY		9505	TCCGCAACTGCGCACTATTGAAAAAATGAAGTAAATTTAATAATGCCANTCCGGAATTT	9564
Dd		2101	TCCGCAACTGCGCACTATTGAAAAAATGAAGTAAATTTAATAATGCCANTCCGGAATTT	2160
OY		9565	AAAAATATATCTGGGAAGATTTGSAAGGGGAAAGCTAATTAABAAAAGTCCGGTAGATCTTAAT	9624
Dd		2161	AAAAATATATCTGGGAAGATTTGSAAGGGGAAAGCTAATTAABAAAAGTCCGGTAGATCTTAAT	2220
OY		9625	CAGGTAGACTGGTATACGATTAATAATATACATGCGATTTCTTGAAAAAGCATACAAAAA	9684
Dd		2221	CAGGTAGACTGGTATACGATTAATAATATACATGCGATTTCTTGAAAAAGCATACAAAAA	2280
OY		9685	TTGACATATTCATCAAGAGAGGAGAAAAAGGAAAAAATTA	9726
Dd		2281	TTGACATATTCATCAAGAGAGGAGAAAAAGGAAAAAATTA	2322
 RESULT 5 US-09-841-786-11				
/	Sequence 11, Application US/09841786			
/	Patent No. 6669940			
/	GENERAL INFORMATION:			
/	APPLICANT: NAGARAJA, T. G.			
/	APPLICANT: STEWART, GEORGE C.			
/	APPLICANT: MARIYAMAN, SANJEEV K.			
/	APPLICANT: CHENGAPPAN, M. W.			
/	TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN			
/	TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF			
/	FILE REFERENCE: 30296			
/	CURRENT APPLICATION NUMBER: US/09/841, 786			
/	CURRENT FILING DATE: 2001-04-24			
/	PRIOR APPLICATION NUMBER: 09/558, 257			
/	PRIOR FILING DATE: 2000-04-25			
/	NUMBER OF SEQ ID NOS: 15			
/	SOFTWARE: PatentIn Ver. 2.1			
/	SEQ ID NO 11			
/	LENGTH: 2141			
/	TYPE: DNA			
/	ORGANISM: Fusobacterium necrophorum			
/	US-09-841-786-11			
 Query Match 22.0%; Score 2141; DB 3; Length 2141;				
Best Local Similarity 100.0%; Pred. No. 0;				
Matches 2141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY		3553	GGAGCTTGTGAAGTGTCTTTCTTCAATATATATTAATAAAGAACATCTGCTGTGCTAGAT	3612
Dd		1	GGAGCTTGTGAAGTGTCTTTCTTCAATATATATTAATAAAGAACATCTGCTGTGCTAGAT	60
OY		3613	GGAGTAAGATTAATTTTAAGGAGCAATAAATAAGTAGAGGTGACTTCTTCTGATTTCT	3672
Dd		61	GGAGTAAGATTAATTTTAAGGAGCAATAAATAAGTAGAGGTGACTTCTTCTGATTTCT	120

QY 3673 ACTTTGTGAGCATGGGGCGGATCTGTCGACTTCAGTGGATCATATTGGAAGTGA 3732
 Db 121 ACTTTGTGAGCATGGGGCGGATCTGTCGACTTCAGTGGATCATATTGGAAGTGA 180
 QY 3733 AATAGCAATCATAGTGTGTTTACGTGAGCGGCTGCTGTAATTAATTCAAGTAA 3792
 Db 181 AATAGCAATCATAGTGTGTTTACGTGAGCGGCTGCTGTAATTAATTCAAGTAA 240
 QY 3793 ACAAGTCTTGTGTTTAAATAGTGAATTGCAAAATGCCAATTAATTAAATGAT 3852
 Db 241 ACAAGTCTTGTGTTTAAATAGTGAATTGCAAAATGCCAATTAATTAAATGAT 300
 QY 3853 TTGAGTGAAGAACTCAAGTACAGCAGAGCAGGTTTGGAAAGCAGTTAAAGAAAGTGA 3912
 Db 301 TTGAGTGAAGAACTCAAGTACAGCAGAGCAGGTTTGGAAAGCAGTTAAAGAAAGTGA 360
 QY 3913 GGAACAAGAAAAAGTTATCTATTGGGAACCTTGCTTCTATCACTTAGTGAACATGA 3972
 Db 361 GGAACAAGAAAAAGTTATCTATTGGGAACCTTGCTTCTATCACTTAGTGAACATGA 420
 QY 3973 GTTTCGCAAAATCGAAAAATAACAGTACAGAGAGAACTGGAAGCCAAAAATGGAT 4032
 Db 421 GTTTCGCAAAATCGAAAAATAACAGTACAGAGAGAACTGGAAGCCAAAAATGGAT 480
 QY 4033 GTTGAATGTCATGCTTATCAAGCGGACACCCAAAGTACAGAGCTTTAAATTACAGCT 4092
 Db 481 GTTGAATGTCATGCTTATCAAGCGGACACCCAAAGTACAGAGCTTTAAATTACAGCT 540
 QY 4093 GGAAGATCAATGGAACCTGTAAGGCGTACTGTAAGCTGTTGCCAATTAACAACAAGTA 4152
 Db 541 GGAAGATCAATGGAACCTGTAAGGCGTACTGTAAGCTGTTGCCAATTAACAACAAGTA 600
 QY 4153 AATGCTTCTATTAGTGGTGGAGATPACTAAAGTTAATCGAGCGAGCGCAAAAGCTCT 4212
 Db 601 AATGCTTCTATTAGTGGTGGAGATPACTAAAGTTAATCGAGCGAGCGCAAAAGCTCT 660
 QY 4213 TTAGCAACCACTCAAGTGACTGCTGAGTGAAGCAGCGAGGAGCAATTAGTTCTGAGAG 4272
 Db 661 TTAGCAACCACTCAAGTGACTGCTGAGTGAAGCAGCGAGGAGCAATTAGTTCTGAGAG 720
 QY 4273 GGAATTAAGAAATTAACAAGGCGCTTTTGTGCAATTAAGATTGACATGACGTGAAGCT 4332
 Db 721 GGAATTAAGAAATTAACAAGGCGCTTTTGTGCAATTAAGATTGACATGACGTGAAGCT 780
 QY 4333 AGCGTTGATTAATCTTCCATGCAAGAGCTAATGAATGATGTCACCAAGATGTC 4392
 Db 781 AGCGTTGATTAATCTTCCATGCAAGAGCTAATGAATGATGTCACCAAGATGTC 840
 QY 4393 AAAGGAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGAAGAAATAAAA 4452
 Db 841 AAAGGAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGAAGAAATAAAA 900
 QY 4453 TATTTAAGAAAGATCGTGTGTTTAATACGACTGGAATGTTATTATACGAAGAACACTA 4512
 Db 901 TATTTAAGAAAGATCGTGTGTTTAATACGACTGGAATGTTATTATACGAAGAACACTA 960
 QY 4513 GAAAAAGCAAAAGAAAAAGAGCGGCTGATGTAATGCTGCTTAAATCCGTTGCTGGA 4572
 Db 961 GAAAAAGCAAAAGAAAAAGAGCGGCTGATGTAATGCTGCTTAAATCCGTTGCTGGA 1020
 QY 4573 ACGGATTAATCCGCTGAGAGAGTATGCAAGTCAATGCTTTAAATTAATTTAA 4632
 Db 1021 ACGGATTAATCCGCTGAGAGAGTATGCAAGTCAATGCTTTAAATTAATTTAA 1080
 QY 4633 GCAGAATTAGTGAAGCAATTAAGAACCGGAGAGGATTAATTAATTCATGCAACATGTA 4692
 Db 1081 GCAGAATTAGTGAAGCAATTAAGAACCGGAGAGGATTAATTAATTCATGCAACATGTA 1140
 QY 4693 AATGTAAGAGCAAAATCATCTACTGTTGTGTAATGCGGCTTCTGAGCTTCTATCAGC 4752
 Db 1141 AATGTAAGAGCAAAATCATCTACTGTTGTGTAATGCGGCTTCTGAGCTTCTATCAGC 1200

QY 4753 AAGATGCTTTTTCAGGAATGGATCTGAGCATGAGCAAGACTTATCAATGACACGATT 4812
 Db 1201 AAGATGCTTTTTCAGGAATGGATCTGAGCATGAGCAAGACTTATCAATGACACGATT 1260
 QY 4813 GCAAGGTGATTAAGGAAGAAATTTCTGCTGATTTCTTAAATGGAACGCAAAATATTC 4872
 Db 1261 GCAAGGTGATTAAGGAAGAAATTTCTGCTGATTTCTTAAATGGAACGCAAAATATTC 1320
 QY 4873 ATTCTGGGAGTAAATGTTTGGCGGAACATTGCCGGTTCCTTTCTACGGCGGTAGAGCT 4932
 Db 1321 ATTCTGGGAGTAAATGTTTGGCGGAACATTGCCGGTTCCTTTCTACGGCGGTAGAGCT 1380
 QY 4933 GCTTTGCAATTAATCTTTCATTAATTAACCTCTGCTTGTGATTAACGAACGAAGTA 4992
 Db 1381 GCTTTGCAATTAATCTTTCATTAATTAACCTCTGCTTGTGATTAACGAACGAAGTA 1440
 QY 4993 AATCCTTTTATGGAAGAAATACAAAGTCAATGTAACAGTTGGAATGATTCTCATATT 5052
 Db 1441 AATCCTTTTATGGAAGAAATACAAAGTCAATGTAACAGTTGGAATGATTCTCATATT 1500
 QY 5053 ACAACGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGGAATCGAGGAATGTA 5112
 Db 1501 ACAACGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGGAATCGAGGAATGTA 1560
 QY 5113 TCTGTCAATCGTGTTCATGTAAGAACGGAACCTTAGTATGATTCGAGTTGAAAGA 5172
 Db 1561 TCTGTCAATCGTGTTCATGTAAGAACGGAACCTTAGTATGATTCGAGTTGAAAGA 1620
 QY 5173 GTTAAGTTCTTTCATGTAAGTACAAAGATTAATTAATTAATTAATTAATTAATTAAT 5232
 Db 1621 GTTAAGTTCTTTCATGTAAGTACAAAGATTAATTAATTAATTAATTAATTAATTAAT 1680
 QY 5233 GCAATGAGGAAAAAGCGCTGAGGTTGGAAGCAACAGTGTCTCATCAATTAATGAAAA 5292
 Db 1681 GCAATGAGGAAAAAGCGCTGAGGTTGGAAGCAACAGTGTCTCATCAATTAATGAAAA 1740
 QY 5293 CAATCAGTTAATGCTATTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5352
 Db 1741 CAATCAGTTAATGCTATTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
 QY 5353 AAAAAATATCAATGTAAGTACGCAAAAGATTAATTAATTAATTAATTAATTAATTAAT 5412
 Db 1801 AAAAAATATCAATGTAAGTACGCAAAAGATTAATTAATTAATTAATTAATTAATTAAT 1860
 QY 5413 GGAAGAGCAAAAGAGCGCTGTCGAAGAGCTTCTGCAAGTACTACCTTAATTAAGCA 5472
 Db 1861 GGAAGAGCAAAAGAGCGCTGTCGAAGAGCTTCTGCAAGTACTACCTTAATTAAGCA 1920
 QY 5473 GTTCTTCTCATGTTGATCAATCTGATTAATGACAAAGATTTAGAGAAAGAAATATGGA 5532
 Db 1921 GTTCTTCTCATGTTGATCAATCTGATTAATGACAAAGATTTAGAGAAAGAAATATGGA 1980
 QY 5533 AATTAAGAAAAAGCAATGTTTAATGTTCTAGTGAATTAAGAGTCAAGTGTCACAAT 5592
 Db 1981 AATTAAGAAAAAGCAATGTTTAATGTTCTAGTGAATTAAGAGTCAAGTGTCACAAT 2040
 QY 5593 GCGACAGTCTTCCGAGAGCAAGTGAACAAGCTGCAAGAGGCTGGAAGTGAAGTAAAT 5652
 Db 2041 GCGACAGTCTTCCGAGAGCAAGTGAACAAGCTGCAAGAGGCTGGAAGTGAAGTAAAT 2100
 QY 5653 AAAATTAACAAAAATCTTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 5693
 Db 2101 AAAATTAACAAAAATCTTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAAT

RESULT 6
 US-09-841-786-12
 ; Sequence 12, Application US/09841786
 ; Patent No. 6669940
 ; GENERAL INFORMATION:
 ; APPLICANT: NAGARAJA, T. G.
 ; APPLICANT: STEWART, GEORGE C.
 ; APPLICANT: NARAYANAN, SANJEEV K.

APPLICANT: CHENGAPRA, M. M.
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
FILE REFERENCE: 30296
CURRENT APPLICATION NUMBER: US/09/841,786
PRIOR APPLICATION NUMBER: 09/558,257
PRIOR FILING DATE: 2000-04-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 12
LENGTH: 1887
TYPE: DNA
ORGANISM: Fusobacterium necrophorum
US-09-841-786-12

Query Match 19.4%; Score 1887; DB 3; Length 1887;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5624 CTGCAATGAGAGCTGAGTAGAGCTTAATTAATTAACCAAAATCTTCTGACATTA 5683
DB 1 CTGCAATGAGAGCTGAGTAGAGCTTAATTAATTAACCAAAATCTTCTGACATTA 60
QY 5684 AAAATGTAATCTCAAAATGTAAGAAATGCTTTGGTAAAGCAAAATCTCATCTTA 5743
DB 61 AAAATGTAATCTCAAAATGTAAGAAATGCTTTGGTAAAGCAAAATCTCATCTTA 120
QY 5744 AAAATGTAATGTAAGAGCTGAGTGGAGCTGAGAGCTGAGAGCTGAGAGCTTCTGAG 5803
DB 121 AAAATGTAATGTAAGAGCTGAGTGGAGCTGAGAGCTGAGAGCTGAGAGCTTCTGAG 180
QY 5804 CAGTGAATGAGATTTGTAATTAATAGATAGCAAAATTAATCATCAAAATCACTGCGA 5863
DB 181 CAGTGAATGAGATTTGTAATTAATAGATAGCAAAATTAATCATCAAAATCACTGCGA 240
QY 5864 AGGGAATGTCGAGATTTATACAGAGCTGATGCGGTAAATGCTTAATTAAGCAAGAA 5923
DB 241 AGGGAATGTCGAGATTTATACAGAGCTGATGCGGTAAATGCTTAATTAAGCAAGAA 300
QY 5924 TGTCTGAGTGGCCCGTGCAGCAATAGAGCCTCAACCAAGTGAATGAATTAACAGAT 5983
DB 301 TGTCTGAGTGGCCCGTGCAGCAATAGAGCCTCAACCAAGTGAATGAATTAACAGAT 360
QY 5984 CTACAAAAGCATATGTAAAGATTTCTACAGATTTGCTTAAGAGAAACAGATGATTA 6043
DB 361 CTACAAAAGCATATGTAAAGATTTCTACAGATTTGCTTAAGAGAAACAGATGATTA 420
QY 6044 TTACTATCTCAAGGCAAGTAAAGTGTAGATTAAGTATTCAAAATCTTTAATTA 6103
DB 421 TTACTATCTCAAGGCAAGTAAAGTGTAGATTAAGTATTCAAAATCTTTAATTA 480
QY 6104 AGGAAGCTTATCAAGAAAAGAAAATAGTAATTAAGAAAGATTTTACCAATAGTT 6163
DB 481 AGGAAGCTTATCAAGAAAAGAAAATAGTAATTAAGAAAGATTTTACCAATAGTT 540
QY 6164 CAGCTACTCATCTTTAAATCTTTATTTGCAAAATGCGCTGGTTCAAGAACAGCCGAG 6223
DB 541 CAGCTACTCATCTTTAAATCTTTATTTGCAAAATGCGCTGGTTCAAGAACAGCCGAG 600
QY 6224 TGGCAGGAATCTGTAATATCAACAGGTTTATGAGAAAACAGAGCTTTGTAAGAAAT 6283
DB 601 TGGCAGGAATCTGTAATATCAACAGGTTTATGAGAAAACAGAGCTTTGTAAGAAAT 660
QY 6284 CTATATTAATGCAAAATTAATCTGTAATTAATCAAGAGTTTACAGAAATCAATGAG 6343
DB 661 CTATATTAATGCAAAATTAATCTGTAATTAATCAAGAGTTTACAGAAATCAATGAG 720
QY 6344 TAGTAGTTCTGTTGGTGTGTGTAATGTAAGAGTAGAGCTTTCTGATACCAATA 6403
DB 721 TAGTAGTTCTGTTGGTGTGTGTAATGTAAGAGTAGAGCTTTCTGATACCAATA 780
QY 6404 TTATTAAGAAATACCAAGACAGAGTTGAAAAATCTACATGTCGATGAAGTTTCG 6463

DB 781 TTATTAAGAAATACCAAGACAGAGTTGAAAAATCTACATGTCGATGAAGTTTCG 840
QY 6464 GAGAAAGCTGAAATTAACAGAGATTTTAAGCAAGAAATTTCTCTTTTGAAGTCGAG 6523
DB 841 GAGAAAGCTGAAATTAACAGAGATTTTAAGCAAGAAATTTCTCTTTTGAAGTCGAG 900
QY 6524 TCGAGAGCCGGGTGAGAGCCGAGAGTGGAGAGACGTTTCCGTAAATCAATTTGAG 6583
DB 901 TCGAGAGCCGGGTGAGAGCCGAGAGTGGAGAGACGTTTCCGTAAATCAATTTGAG 960
QY 6584 GAAAGCAGAGATGATGTGGAAGCAAGATTTTGGTAAAAAAGCTGAGATTACAG 6643
DB 961 GAAAGCAGAGATGATGTGGAAGCAAGATTTTGGTAAAAAAGCTGAGATTACAG 1020
QY 6644 CAAAACGTTATAGTTCTGTTGCAATTTGAAATGCCCGAGTGGAGTGGCTCAAAAGAG 6703
DB 1021 CAAAACGTTATAGTTCTGTTGCAATTTGAAATGCCCGAGTGGAGTGGCTCAAAAGAG 1080
QY 6704 CTGGAATGAGAGCAGAGCTGAGTACCAAGATGAATCAAAACAGAGAGAGTGA 6763
DB 1081 CTGGAATGAGAGCAGAGCTGAGTACCAAGATGAATCAAAACAGAGAGAGTGA 1140
QY 6764 AAAATTTAAATTAATGACTCGAAAACAGATTAGATGTAATAGCAAGAAATGAGATTAAT 6823
DB 1141 AAAATTTAAATTAATGACTCGAAAACAGATTAGATGTAATAGCAAGAAATGAGATTAAT 1200
QY 6824 CAGTACTGGAATGCGTTACCGCGAGCTGGAATTTCTGACCGGAGATTTGAGAGTGG 6883
DB 1201 CAGTACTGGAATGCGTTACCGCGAGCTGGAATTTCTGACCGGAGATTTGAGAGTGG 1260
QY 6884 TTTCTGTCAATTAATTTGCCAATTAAGTAAGAAACAGATGATGATGATCTTACACT 6943
DB 1261 TTTCTGTCAATTAATTTGCCAATTAAGTAAGAAACAGATGATGATGATCTTACACT 1320
QY 6944 CTTTACTGATGTAATTTAAAGCTCTTAATTAATTTGCAATTTCTTGAACGCGGTG 7003
DB 1321 CTTTACTGATGTAATTTAAAGCTCTTAATTAATTTGCAATTTCTTGAACGCGGTG 1380
QY 7004 GAGAGCCGAGAGCTTTCAGAGCAATTAACCGAGTGGTTCTGTTAACTATTAATGTT 7063
DB 1381 GAGAGCCGAGAGCTTTCAGAGCAATTAACCGAGTGGTTCTGTTAACTATTAATGTT 1440
QY 7064 CTGTATAGCTCGAGTTCAATTAATCTGATTTTGAATTTCTGATGAGAGAAAGTAAAG 7123
DB 1441 CTGTATAGCTCGAGTTCAATTAATCTGATTTTGAATTTCTGATGAGAGAAAGTAAAG 1500
QY 7124 TAAAGCAAAAGAGAAAAAATATTAAGCAAAACAGCAAAATCAGGAATCGAGAG 7183
DB 1501 TAAAGCAAAAGAGAAAAAATATTAAGCAAAACAGCAAAATCAGGAATCGAGAG 1560
QY 7184 CAGCAATCGAGCAATGCTTTGTAATTAATTTTGAACAGCTGTAGAAATAGAAAA 7243
DB 1561 CAGCAATCGAGCAATGCTTTGTAATTAATTTTGAACAGCTGTAGAAATAGAAAA 1620
QY 7244 ATTTGAGAGAAAGAAACAGAGTTTAAATCTTTGAAGAGTTAACAAAGAACAG 7303
DB 1621 ATTTGAGAGAAAGAAACAGAGTTTAAATCTTTGAAGAGTTAACAAAGAACAG 1680
QY 7304 ATTAAGAAATGAATGATGCTTACGAAATAATCTTCAATCAGCAGTATTTCTACAGAG 7363
DB 1681 ATTAAGAAATGAATGATGCTTACGAAATAATCTTCAATCAGCAGTATTTCTACAGAG 1740
QY 7364 ATACTTCTGTAAGCCGATAGAGAGATTAACAGGAGAGAAATTAAGCCATTGTA 7423
DB 1741 ATACTTCTGTAAGCCGATAGAGAGATTAACAGGAGAGAAATTAAGCCATTGTA 1800
QY 7424 AGACTTTCTGATTTATTTGAAAAATGTAATTAATCAACAGAGCAAGAAATTAATCA 7483
DB 1801 AGACTTTCTGATTTATTTGAAAAATGTAATTAATCAACAGAGCAAGAAATTAATCA 1860
QY 7484 CTTTACTGTTGTTGGAACTGACAG 7510

Db 1861 CTTCTACTGCTGGTTGGAACTGCAG 1887

RESULT 7
US-09-841-786-9
Sequence 9, Application US/09841786
Patent No. 6669940
GENERAL INFORMATION:
APPLICANT: NAGARAJA, T. G.
APPLICANT: STEWART, GEORGE C.
APPLICANT: NARAYANAN, SANGEER K.
APPLICANT: CHENGAPPA, M. M.
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
FILE REFERENCE: 30296
CURRENT APPLICATION NUMBER: US/09/841,786
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 09/558,257
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 1130
TYPE: DNA
ORGANISM: Fusobacterium necrophorum
US-09-841-786-9

Query Match 11.6%; Score 1130; DB 3; Length 1130;
Best Local Similarity 100.0%; Pred. No. 9,78-221;
Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACGGCATCAAAATAATACGTTGAGAGACAAAGAGATATCAGATTCTAAAAA 60
Db 1 ATGACGGCATCAAAATAATACGTTGAGAGACAAAGAGATATCAGATTCTAAAAA 60

Qy 61 GTTTTAATGATTTGGATGTTGATTAACACTATGACGGTGAAGGCTTAATGATCAATC 120
Db 61 GTTTTAATGATTTGGATGTTGATTAACACTATGACGGTGAAGGCTTAATGATCAATC 120

Qy 121 ACCGGAGTGAATTTTGGACAAATAATAGAAAAAGATTAAGTTTATGACTTACT 180
Db 121 ACCGGAGTGAATTTTGGACAAATAATAGAAAAAGATTAAGTTTATGACTTACT 180

Qy 181 ACAACAGATTTCAAGGAGAGACGCTTTTACAGTTTATGATTTGCTTTACAGAA 240
Db 181 ACAACAGATTTCAAGGAGAGACGCTTTTACAGTTTATGATTTGCTTTACAGAA 240

Qy 241 AATAATATAGCAATCTATATTTTGGGAAAAAGAAATAGACGGGGTAAATATCTTTT 300
Db 241 AATAATATAGCAATCTATATTTTGGGAAAAAGAAATAGACGGGGTAAATATCTTTT 300

Qy 301 AACTTGTCAATGAAAAATGAGTAGATGAGTATTCACACGGAATTCGAAAAATATAA 360
Db 301 AACTTGTCAATGAAAAATGAGTAGATGAGTATTCACACGGAATTCGAAAAATATAA 360

Qy 361 ATTGAGAGAAATTTATTTCTTAAGCTCGAAGGAGATGCGAGTGAAGAAAAATGAGTT 420
Db 361 ATTGAGAGAAATTTATTTCTTAAGCTCGAAGGAGATGCGAGTGAAGAAAAATGAGTT 420

Qy 421 ATCAATGCTGGTCTTTTCAATCTATATTCGAAAAACAAGATGATTTTAAAGAGCTTTG 480
Db 421 ATCAATGCTGGTCTTTTCAATCTATATTCGAAAAACAAGATGATTTTAAAGAGCTTTG 480

Qy 481 GAAAGAGCCAAACATGCTAAAGTTTAAATGAAATCATTCGATGATGGAAGATTAATAA 540
Db 481 GAAAGAGCCAAACATGCTAAAGTTTAAATGAAATCATTCGATGATGGAAGATTAATAA 540

Qy 541 ATTCCATGATTCGAAATGAGAGCATTAAGTGAAGAAAAATCAATGCTTTGAAGGC 600
Db 541 ATTCCATGATTCGAAATGAGAGCATTAAGTGAAGAAAAATCAATGCTTTGAAGGC 600

Qy 601 ATCCGTTTATATCGCGGATTTAGATGGAAGATGACTGCAATCTAAGAAGAGAAAT 660
Db 601 ATCCGTTTATATCGCGGATTTAGATGGAAGATGACTGCAATCTAAGAAGAGAAAT 660

Db 601 ATCCGTTTATATCGCGGATTTAGATGGAAGATGACTGCAATCTAAGAAGAGAAAT 660
Qy 661 ACAGATTTTAAATTTAGTCAATATTAGTGAATTAATTTGCTGCTGACCGGAGAT 720
Db 661 ACAGATTTTAAATTTAGTCAATATTAGTGAATTAATTTGCTGCTGACCGGAGAT 720

Qy 721 TTAAGAGTCAACAGACAAATCTGAGATATTAATTTTCTGAGCTCACATGATTTCTCT 780
Db 721 TTAAGAGTCAACAGACAAATCTGAGATATTAATTTTCTGAGCTCACATGATTTCTCT 780

Qy 781 CAAAAAGCTATGGAAAAAATTTCACTGTTGGAAAAGAGATTAAGAAATATGTAAGGA 840
Db 781 CAAAAAGCTATGGAAAAAATTTCACTGTTGGAAAAGAGATTAAGAAATATGTAAGGA 840

Qy 841 AATCCAAAGCAATATTTGAATCTGATGCTGATTTGGAAGAGATGGAATATATAAAT 900
Db 841 AATCCAAAGCAATATTTGAATCTGATGCTGATTTGGAAGAGATGGAATATATAAAT 900

Qy 901 AGTGGAAAGCTACAAATGGAGATTTATTAAGAAAGAGGAAAAAGAACTTATAC 960
Db 901 AGTGGAAAGCTACAAATGGAGATTTATTAAGAAAGAGGAAAAAGAACTTATAC 960

Qy 961 ACTCCTTAAGTTATACAGATGTTGAAGCTTCCTTAAGATTAATTAAGAAAAAGTCA 1020
Db 961 ACTCCTTAAGTTATACAGATGTTGAAGCTTCCTTAAGATTAATTAAGAAAAAGTCA 1020

Qy 1021 GAAAGAAATGTTGACATTAAGTGAAGCAAGAAATTTCTATGATGCAACTTTAGTACT 1080
Db 1021 GAAAGAAATGTTGACATTAAGTGAAGCAAGAAATTTCTATGATGCAACTTTAGTACT 1080

Qy 1081 AAGCTTGAAGCACTCTTTAGCTTTGTTACAGTTTCTATTTCTCTAT 1130
Db 1081 AAGCTTGAAGCACTCTTTAGCTTTGTTACAGTTTCTATTTCTCTAT 1130

RESULT 8
US-08-973-462-2
Sequence 2, Application US/08973462B
Patent No. 6191270
GENERAL INFORMATION:
APPLICANT: DRULHE, PIERRE
APPLICANT: DAUBERSIES, PIERRE
TITLE OF INVENTION: MALAIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
FILE REFERENCE: 0660-0125-0 PCT
CURRENT APPLICATION NUMBER: US/08/973,462B
EARLIER FILING DATE: 1998-02-06
EARLIER APPLICATION NUMBER: PCT/FR96/00894
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: FR 95/07007
EARLIER FILING DATE: 1995-06-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 5361
TYPE: DNA
ORGANISM: P. falciparum
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(5361)
US-08-973-462-2

Query Match 1.2%; Score 116.2; DB 3; Length 5361;
Best Local Similarity 41.7%; Pred. No. 8,4e-14;
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

Qy 5126 GTTCGATGAACCGAAGCTTTAGTATGATTCGAGTTGAAGAGAGTAAATTTCTTCA 5185
Db 773 GTTCGATGAACCGAAGCTTTAGTATGATTCGAGTTGAAGAGAGTAAATTTCTTCA 5185

Qy 5186 ATGTGATGCAAAAGATTAATAATTAATCA---TTGCCGAAATGCAATGAG 5242
Db 833 ATGTGATGCAAAAGATTAATAATTAATCA---TTGCCGAAATGCAATGAG 5242

Db 3044 AAGGATACAGAAAAATTATTACAGGTANG-----TTTCGAAGTATAGAACCACTA 3097
Qy 7451 TAGATATTACACAGAGACAAGAAATTAATCATCTTACTGCTGTTGGAACTGCAG 7510
Db 3098 TAGATATCCATCAGAGAAAAAGTTGATTTGATGAAAAATGTTGTTGATTTTAA 3157
Qy 7511 GTCTTGCTCCGATCAGAACAGTGGCAGTTACAATATTAAAAAGAAATTCGGAGTTA 7570
Db 3158 ATTAATATGAAAAATATGAAAGAGGTTTATTAATTAATTAAGAAATATTTCAAGTACTG 3217
Qy 7571 CTGTGAAAAATCTTTTGAAAAAGCAGCTGAAAAAGTAAATCTTATGATCCGATAT----- 7625
Db 3218 AAGGTTCCAGAAAACTGTAACTGAACATGTAACAAATGTATATGTGATGTTGATG 3277
Qy 7626 ----TACAGAAATGTGCTTTAACAGCATATCAAGTCTGTAGAGCATTTGGAGATAG 7681
Db 3278 TTCCGTCTATGAAATCAATTTTAAAGAAATTAAATGAGCAGAGAGGTTGAAAGAAA 3337
Qy 7682 GAGCTGCTATGCAAAATTAATTTCTAAATGGAAGTCAAAATATCAAGTATTAATAATTTCTA 7741
Db 3338 TGTTTTTAAATTTGGAAGATGATTTTAAAGTGAAGTATGTAATTTACTGTAGAAAGAA 3397
Qy 7742 AGCTATTAGAAAAAAATATGATGTTATTGTAAAGATTAATTCGAATTTGAGCGAGAG 7801
Db 3398 TTAAGGATGAACCGGTTCAAAAAAGAGTGAAGAAAAAGAACTGTAGTATTAATGAAAGAA 3457
Qy 7802 CAAAAGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTAATCTCAAAAGCAAGAAATG 7861
Db 3458 TGGAGAAAAATTTGATGATGATTTAGAGAGAAAAAGAAAGATTTAAACAGCAAGATGA 3517
Qy 7862 AAATGAATTCAGAGGTTGAAATTTGAAGAGATTAATTTCAATGAGAAAAATAGACTA 7921
Db 3518 TAGATGCAAGTAAATCCATAGAAATATCTTCAGATTTCTAAAGAAAGAACTGAATCTA 3577
Qy 7922 GCCCTTAAAGAAATTTGGAAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGTA 7981
Db 3578 TTAAGATTAAGAAAAAGATGTTTCACTAGTGTGAAGAAATTCGAACATGATATAG 3637
Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGCAGTAGCAGGGCAG 8027
Db 3638 ATGAAGTGTGAGAAAGTTTNGAATTTGAAAAATATGGAAGAGAGTTAATGAAGATG 3697
Qy 8028 AATTATTTCCGAAGCAAAAGATGCGGAAGCTTTATTTGAAAGTTAGTACAAAATCCGG 8087
Db 3698 CTGTTGAATTAATGACATTACTAGCAAACTTATGGAAGAACTCAAGAGTTAAATGAAG 3757
Qy 8088 AAGAAGTATTTTCTATGCAGATATGTAATTTGAGACACATTAATGAAGTAAC 8147
Db 3758 TAGAAGCAGATTTTAAAGATATGAAAGAAATTAAGAAATTAAGAAAGC--ATTATC 3815
Qy 8148 AGCAGTTTCTAAAGCAGTAACAGGTTCTGTATTTGGAGAGAGTTGAGTCACCAAGCAGTA 8207
Db 3816 AGAAGATTTTAAGAAATTAATGATGCAAAAGATATACATTAGAAAAAGTTATGAAAGA 3875
Qy 8208 AGTACTGCTGAGGTAAAACTATGTAAGTAAGTAAGAAAGAAATTTGTCAGAACAA 8267
Db 3876 GGAACATGATTAACGACGAGCTGGATGAGAGTTGTAGAAATTAAGAAAGTGAAGAGA 3935
Qy 8268 TCGATTAATGCAATTTCTAAAGTGAAGGTTTGAATGAAGTAAGTAAGTCTGTAATC 8327
Db 3936 CAAGATCGAAAAAGATCTGATTTAAAGATTTTGAAGAAAGTATTAATTAAG----- 3988
Qy 8328 TTCTGTAGTATCAGAAAAATGAGAGAAATTTGCCGAGCAGAGAGTAACTTTACAGC 8387
Db 3989 --AAATTAAGAAATCAAGAACTTGAAGTGAATTTTGAAGATTATTAAGAAATTTAA 4046
Qy 8388 ACAAGTATATCTGAATCCGATGTTGTTAGAAAGCAAGTTATGAAGTAATGATTA 8447
Db 4047 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATAGAAAAAGATCATTTTGAAGAA 4106
Qy 8448 CACAAAAAAATATATTTTCAAGAGTCAATGCTCTTGCTTTAAATGATACAAAGATGAAGC 8507
Db 4107 ATTGAGAGAAAGCTGAAGAAATTAAGATCTTGAAGCAGATATATTAATTAAGAAAGTATC 4166

Qy 8508 GAATATGAATCTTAGCGGTAGCCGGTGTGCATGCAACAAGAACAAAGCATTTAC 8567
Db 4167 TTCAATTGAAGTTGAAGAGAAAAAATTGAAGAAAGTACACGAATTAAGAAAGAGAGT 4226
Qy 8568 GAGTCAAAAGATTAACTTTACACTGTAATGAGAGAAAGTATCTCACTTCGTGC 8627
Db 4227 AGAATATATTAATGATGATGTCGCAATATTAAGGTTTGAAGAAAGTATTAAGAA 4286
Qy 8628 AAAAGCTTTGGCTAAAAATGAAAAATTATGGAATGTAAAGAACTGAGAGAGCTTAGT 8687
Db 4287 AGTGAATGA--TTTAAAGAAAGTATATTAAGCATGTTAAAGGAGATATGGA--ATTAGG 4343
Qy 8688 CGGACCGGAAACACAGCCGTGAAAAATTATACAAAGAGTACTACAGAGCATTTGTTGC 8747
Db 4344 GGATATGATTAAGAAAGTTTAGAAGATGTAACAACAAACTTGAAGAAAGATGATC 4403
Qy 8748 AGAATAATTGGAAAAATTGAGATTAATTAAGAAACATTCGAAGATTAATGATTTGTAAG 8807
Db 4404 CTTAAAGATGTTTATCTAGTCAATTAAGCATGATGTAAGAACAAATGAAACCAAGAAA 4463
Qy 8808 AGTCAACGAGACGGAACCAAGAGCTCTGTGGAAGAAATGTAATTTCTGTGAAGAAA 8867
Db 4464 AAAAGCTCAAGACCTAAGTTGGAAGAGTATTTTAAAGAGAGAGTTAAAGAAAGACC 4523
Qy 8868 TACAATTTCAAGGGGAAACAAATCATCCATTGAAGATTAAGCCAGAAATTTGGAACCCG 8927
Db 4524 AAAAGAAAAAATTAACAAAAAGAAAGTAAAGTTGATTAATTAAGATTAAGAACCAAGAA 4583
Qy 8928 AAGTGAATGTAGATGCTTTGAAATGAACCTTGAATGATTTCAAGAAAAAGTGTGG 8987
Db 4584 TGAATATGTAAAGTTGAAATGAAGAAATGAAGATATAGAAAGATGTAGAAAGATAT 4643
Qy 8988 CTATGTGGAATTGGAATTTGGAATTTGATGATTAATTAATGATTAAGAAAAATGTGA 9047
Db 4644 AGAAGAAATATAGAAAGAGTAAAGTTAGAGATTAAGTAAGATATAGATGAAGATAT 4703
Qy 9048 AGCCAAATCGAAGACATGCTATTTAGAAATCACTGGAAGAAACAGAAATTAACAACATT 9107
Db 4704 AGGTGAAGACAAAGATGAGATTATTAATTAAGTCCAAAAAGAGAAAGCATTTGAAAA 4763
Qy 9108 TACAAGCAAAAGTAAATATTTCTGGAAGAAAGAGCGCTGACAGCTGCATATTC 9167
Db 4764 GGTTAAGCGAAAAAGAAAAATTAAGAAAAAGTTGAAGAAAGTGTAGTGTCTTAA 4823
Qy 9168 GAATGTACACTTTCCATGATGATTAATTAATAATTTGCCAAGCAGTATGATCTTC 9227
Db 4824 AAAACAGTAGACAGAGTAAATGAAATATGTTCAAAAAATGATTAAGAAAGTTGAATAAGA 4883
Qy 9228 TCAATTATAACCAAAAAATTCAAAAAATTAATTAATTTAGTATCAATCAA 9274
Db 4884 AGTATCTAAAGCTTTAGATCAAAAAATGATTAATGTTTAA 4930

RESULT 9
US-08-973-462-1
; Sequence 1, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRULHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALALIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973.462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6152

TYPE: DNA
; ORGANISM: P. falciparum
US-08-973-462-1

Query Match 1.2%; Score 116.2; DB 3; Length 6152;
Best Local Similarity 41.7%; Pred. No. 8.8e-14;
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

QY 5126 GTTCGATGAAACGAGAGCTTTAGTTAGATTCTGAGTTGAAGAGTAAGTCTTTCA 5185
DB 1017 GTGTTGAAGAAAGTAGCTTCAAGTGTGATGAAGATATGATTCAAGTATTTGAAGAA 1076
QY 5186 ATGTGATGCAAAAGATCAAAAACAATAATACAA---TTGCCGGAATGCAATGAG 5242
DB 1077 ATGTAGCTCAACTGTGAAAGAAATGTAGCTCCAAGTGTGTAGAAATGTGCTCCA 1136
QY 5243 GAAAACCGCTGAGTTGAGCAACAGTTGCTCATACAAATATTGGAATAACATCAATTA 5302
DB 1137 GTGTTGAAGAAAGTAGAGAAAGAAATGTTGAAGAAAGTAGCTGAAATGTTGAAGAA 1196
QY 5303 TAGCTATTGTAAAAACAGTAAATTTACAACGCGCATATCATAGTAGAAAAATATCA 5362
DB 1197 GTGTAGCTGAATGTGGAAGAAAGTAGCTGAAGAAATGTTGAAGAAAGTAGCTGAAA 1256
QY 5363 ATGTGACTGCAAAAGATTATCTATAGCAATATCTATAGCAGTCGAGTTGAGAGCA 5422
DB 1257 ATGTGGAAGAAATCGTAGCTCCAAGTTGAAAGAAATCGTAGCTCCAAGTTGAAAGAA 1316
QY 5423 AAGGAGCTCTGTGCAAGAGCTTCTGCAAGTACTACTTGAATAGACATTTCTTCTC 5482
DB 1317 TTGTAGCTCCAAAGTGTGTAGAAAGTGTGCTCCAAGTGTGAAGAAAGTAGAGAA 1376
QY 5483 ATGTGATCAAACTGATATTGACAAAGATTAGAGAGAAATTAAGAAATAGGAA 5542
DB 1377 ATGTGAAAGAAAGTAGCTGAAATGTTGAAGAAAGTAGCTGAAATGTTGAAGAA 1436
QY 5543 AAGCAAAATGTTAATGTTTCTAGCTGAAATACAGTCAAGTGTCAAAATGGAACATGTC 5602
DB 1437 GTGTAGCTAAATGTGGAAGAAAGTAGCTGAAATGTGGAAGAAAGTAGAGTGA 1496
QY 5603 TTTCGCGAACAGTGAACAGCTGACAGTGAAGCTGAGTAGCAAGTTAATAATTTAC 5662
DB 1497 ATGTGAAAGAAATCGTAGCTCCAAGCTGTGAAAGAAATCGTAGCTCCAAGTTGAAGAA 1556
QY 5663 AAAATCTTCTGACATATPAAAAATAGTATCTCAAAATGTACGAAATGCTTGTGTA 5722
DB 1557 TTGTAGCTCCAAGTGTGTAGAAAGTGTGAGCTCCAAGTGTGGAAGAAAGTAGAGAA 1616
QY 5723 GC---AAATCTCATTTATCTATTAACAATTGGAATTGAGCTGGAAGTTGAGCTGAG 5779
DB 1617 ATGTGAAAGAAAGTAGCTGAAATGTTGAAGAAAGTAGCTGAAATGTTGAAGAA 1676
QY 5780 GAGCTGAGTGAACAGGTTCTGTAGCAGTGAATGAATGTTGAATTAATACATGAGAT 5839
DB 1677 GTGTAGCTGAATGTGGAAGAAAGTAGCTGAAATGTTGAAGAAAGTAGAGTGA 1736
QY 5840 TAAATCATGCAAAATCACTGCGAAGGAAATGTGCGAGTTATTAAGAGTCTGATCGG 5899
DB 1737 ATGTGAAAGAAAGTAGCTGAAATGTTGAAGAAAGTAGCTGAAATGTTGAAGAA 1796
QY 5900 TAAATTTCTAATTATGACAGAAACAGTGTGGAAGTGGCCCTGACAGAAATGAGCTCAA 5959
DB 1797 TCGTAGCTCCAACTGTGGAAGAAATCGTAGCTCCAAGTTGGAAGAAATGTGAGTCCA 1856
QY 5960 CCAAGTGAATGAATTTACAGATCTACAAAGCATATGTAAAGATTTCTACAGTATG 6019
DB 1857 GTGTGTGAAAGTGTGCTCCAAAGTGTGAAGAAAGTAGAGTGAAGAAATGTTGAAGAA 1916
QY 6020 CTAAAGAGAAACAGATGATTATTTATCTACTAAGGCGCAAGTAGAATAAAGTGTAGATA 6079
DB 1917 GTGTAGCTGAATGTGTAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA 1976
QY 6080 AAGTATTCAAAATCTTAAATTTAAGCAAGACTTATCACAAAAAAGAAATAAGTAATA 6139

DB 1977 ATGTGGAAGAAAGTAGTACT-----GAAATGTGGAAGAAATCGTAGCTCCAACTGTG 2030
QY 6140 AAAAAGATTTGTATACCAATAGTCTAGCTACTCATCTTTAAATCTTATTTGCAATG 6199
DB 2031 AAGAAATGTGTAGCTCCAACTGTGAAGAAATTTAGCTCCAAAGTGTGTA--GAAAGTG 2087
QY 6200 CGCTGTTTGAAGCAAGCCGAGTGGCAGGAACTGTATTAATCAACAAGTTTATGAG 6259
DB 2088 TGGCTCAAGTGTGGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATG 2147
QY 6260 AAACAGAACTCTGTGTAAGAAATCTATATTAATTAATGCAAAACATTAATCTGTA 6319
DB 2148 TTGAAGAAAGTAGTACTGAAATGTTGAAGAAAGTAGTACTGAAATGTTGAAGAAATG 2207
QY 6320 GAGATTACAGAAATTCAGATTCGAGATAGTATGTTGTGTGTGTGTGGAATGTAAGAG 6379
DB 2208 TAGCTCCAACTGTGGAAGAAATCGTAGCTCCAAGTTGAAAGAAATGTAGCTCCAAGTG 2267
QY 6380 TAGAGCTTCTGTGATCCAAATATTTATTAAG---AAATACCAAGCAAGAGTTGGA 6436
DB 2268 TTGTGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTG 2327
QY 6437 AAACCAATGTCTGATGAAGGTTTCGAGAGAAAGCTGAATTTACAGCAGATTCTAAGC 6496
DB 2328 TAGCTGAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATG 2387
QY 6497 AAGAAATTTCTCTTTTGAAGTCCGAGATCGACGACCGGGGTAGAGCCGAGTGCAG 6556
DB 2388 TTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATGTAGCTCCAAGTGTGAAGAAATG 2447
QY 6557 GAACCGTTTCGTAAATCAATTTGCAAGAAAGACGGAAGTAGATGGAAGACAAAGA 6616
DB 2448 TAGCTCAACTGTGAAGAAATTTAGCTCCAAGTGTGTAGAAAGTGTGCTCCAAGTG 2507
QY 6617 TTTTGTAAAAAAGCTGAGATTACAGCAAAACGTTATAGTCTGTGCAATGGAATG 6676
DB 2508 TTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTG 2567
QY 6677 CCGCAGTGGAGTGGCTGCAAAAGAGCTGCAATTTGAGACAGCAGTGCAGATTACCAAG 6736
DB 2568 TAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTCAA 2627
QY 6737 ATGAATCAAAACAGAGAGCAAGTGAAGAAATTTCAAAAT--ATGACTCGAAACAACT 6793
DB 2628 TTGAAGAAATGTAGCTCCAAAGTGTGAAGAAAGTGTAGCTCCAAGTGTGAAGAAAGTG 2687
QY 6794 TAGATGATATAGCAAAATGAGATTAATCAGTACTGGAATCGGTTCAAGCCGAGCTG 6853
DB 2688 TTGCTGAAGAGCTTCCAAATTTATCAGACAACTTTTAAGTAAATTAATTAGAGTGTA 2747
QY 6854 GAATCTTGACGCGGAGTATCTGAGAGTGTCTGTCAATATTAATTGCAATTAAGTGA 6913
DB 2748 TCGAAACTGAGGAAATTAAGGACAGATATTAATGAAGATAGAGAAAGTGAAGAAATG 2807
QY 6914 AAACAGATATGATCATATGATTTTACACTCTTTACTGATGTAATGTAATGAAGCTTTA 6973
DB 2808 TAGTCAACCAATACAGAAACGTGAAGAAACATCAAGCTGAAGTGTAACTACTTTTA 2867
QY 6974 ATTAATTTGCAATTTCTTGAACAGCCGCTGAGAGACCCGAGTCTTGAAGCAATTA 7033
DB 2868 GTTACATATTAAGAGATATCAAGAAATATCAATTAATGATTAATTAAGAGAAATAT 2927
QY 7034 GAGTGTCTCTGTAACTATTAATATGTTCTGTATTAAGTCAAGTTCACATTAAGTCTG 7093
DB 2928 TAGAAGAACTCCACGAAATGTATTAAGTCCGCTTTAAGAAATACCAAGAGTGAAGAG 2987
QY 7094 ATTTGACTTCGTAAGAG---AAAGTAATGTAAACGCAAAAGAGAAATTAATTA 7150
DB 2988 AAAAGAAAGATATGATGTAATTTGAAGAAAGTGAAGAAAGTGTGCTACCACTTTAA 3047
QY 7151 ACGAAACAGCAATATGCAAGATTCGAGAGCAGCAATCGAGCCAAATGCTTGTGTA 7210

Db 3048 TAGAACTGTGGACAGCAGAGAAAGCGCAATATCAATTACGAAATATTGGAAA 3107
Qy ATTAATTTTGGACAGCTGTAGAAAGATAGAAAAATTTGGAAGGAAAGGACAGAGTTT 7270
Db 3108 ATTTGAAAGAAATGAGTAGAAAGTATGAAAAATGTTGCAGAGATTTTGAAGAAATTA 3167
Qy 7271 TAAAACTTTAGACAGAGTTTAAACAAAGAACAGATAAAAAGCTAAATGATGCTACGAAA 7330
Db 3168 ACGAAACTGTATTTACTGTATTAGATTAAGTAGAGAAACAGTAGAAATTTACCGAG 3227
Qy 7331 AAATCTTACATACAGAGTATTCTTACAGAAAGATACCTTGTATAAGCGATAGAGAG 7390
Db 3228 AAAAGTTTGAAGAAACAATGAAATGATTAAGCATTTTGTAGTAAATATTGTGTA 3287
Qy 7391 ATACTCAGGAGAGAGAAATTAAGCCATTTGTAAGCATCTGTATTTTGGAAAAATG 7450
Db 3288 AAGGAATACAGAAAAATTTATTAACAGTATG-----TTTGAAGTATGAAACCACTA 3341
Qy 7451 TAGATATTACACAGAGGACAAAGAAATATATCACTTCACTGTGTGGAGCTGCGAG 7510
Db 3342 TAGTATCCAAATCAGAGAAAGAGTTGATTTGAAATGAAATGTGTTAGTTTGA 3401
Qy 7511 GTCTGCTTCCGATCAGAGACAGTGGCAGTTTCAAAATTTAAAGAAATTCGAGTTA 7570
Db 3402 ATAAATATAGAAAAATATGAAAGAGTTTATTAATTAATTAAGAAAAATATTTCAAGTACTG 3461
Qy 7571 CTGTGAAATTTCTTTGTGAAAGCAGCTGAAAAAGTAAATGTTAGTCGAGTAT----- 7625
Db 3462 AAGGTGTTCAAGAACTGTACTGAACATGTAGAACAAATGTATATGTGATGTTGATG 3521
Qy 7625 ----TACAGAAATGTGCTTTTACAGCATATCAAGTCTGTAGAGACATTTGGAAATAG 7681
Db 3522 TTCCGTATATGAAAGATCAATTTTAAAGAAATATTAATGAGCAGAGAGGTTGAAAGAA 3581
Qy 7682 GAGCTGCTTACAGAAATTTAAATTTCTAATGGAAGTCAATATCTGATTTAAAAATTTCTA 7741
Db 3582 TGTTTTTAAATTTGAAAGATGTATTTAAAGTGAATGTATTAATTTACTGTAGAGAA 3641
Qy 7742 AGCTATTAGGAAAAAATATGATGTTATTTGTAAGAAATCGAATCGAATTTGAGAGCGAG 7801
Db 3642 TTTAAGATGACCGGTTCAAAAAAGGTTAGAAAAAGAACTGTATGATTTATTTGAAGAA 3701
Qy 7802 CAAGAGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAGAAATG 7861
Db 3702 TGGAGAAAAATTTGTAGATGTATTAAGAGAGAAAAAGAGATTTTACAGCAAGATGA 3761
Qy 7862 AAATGATTTACAGAGTTGAAATTTGAGAGAGTATTTTCAATGAGAAAAATGAGTACTA 7921
Db 3762 TAGATGCAAGTAGAAATCCATAGAAATATCTTCAGATTTCTAAAGAAAGAACTGATCTA 3821
Qy 7922 GCCCTTCTAAAGGAATTTGGAAGAGAAATCAATGTCAAAGTGGAAAAAGAAACAGAGTGA 7981
Db 3822 TTTAAGATTAAGAAAAAGATGTTTCACTAGTGTGGAAGAGTTTCAAGCAATGATATG 3881
Qy 7982 CTGCTGAATCTCAGAGAGCT-----TCTGTAGAGCAGTAGACAGGGCAGG 8027
Db 3882 ATGAAAGTGTGAGAAAGTTTAAATTGAAAAATATGGAAGAGAGGTTTAAAGAGATG 3941
Qy 8028 AATTATTTCCGAGCAAGATGCGGAAAGCTCTTATTTGAAAGTTAGTACAAAATCCGG 8087
Db 3942 CTGTGAAATTAATATGACATTACTAGCAAACTTATTTGAGAGAACTCAAGAGTTTAAATGAG 4001
Qy 8088 AAGAGTATTTTCTACAGATATATGTAATGGAAGCAACATTAATTAATGAAGTATAC 8147
Db 4002 TAGAAGCAGATTTAATTAAGATATGGAAGAAATTAAGAAATTAAGAAAAAGC--ATTATC 4059
Qy 8148 AGCAGTTTCTAAAGCAGTACAGGTTCTGTATTTGGAGAGGTTGAGTCAACCAAGCAGA 8207
Db 4060 AGAAGTTTCTAAAGAAATTAATGATGCAAAAGATGATATCAATTAGAAAAAGTTATGAGAA 4119
Qy 8208 AGCTACTGTCGACGTTAAATCTATGTAGAGAGTTGAGAGAGAAATTTGTTCAAGACAA 8267
Db 4120 GGAACATGATATTAACGACGATGTGATGAGTTGAGAAATTTAAAGATGTCGAAGAGAA 4179

Qy 8268 TCGATGGAATCAATTTCTAAGTAGAAGGTTTGGATGAAGATATAAGTACTGCTAAATC 8327
Db 4180 CAAGATGAAAAAATATCTGATTTTAAAGATCTTGAAGAGATATATTTAAAG----- 4232
Qy 8328 TTCTGTAGTATCAGAGAAATGAGAGAGAAATTCGAGACAGAGTGAATTTACTTACAGC 8387
Db 4233 --AAGTAAAGAAATCAAAAGAACTTGAAGAGAAATTTTAAAGATTTTAAAGAAATTTAA 4290
Qy 8388 ACAAGTATATCTGAATCCGATGTTGTTTACGAAGCAAGATTAATGAAAATTAATGATTA 8447
Db 4291 AACTTTGAAACAGATATTTTGAAGAGAAAAAGAAATGAATGAATGAATTTGTTGAAA 4350
Qy 8448 CACAAAAAATATATTTTCAAGAGTCAATGCTCTTGTATTAATGATACAAAGAAATGAAGC 8507
Db 4351 ATTGGAAGAGAGACTGAAGAAATTAAGATCTTGAAGCAGATATATTTAAAGAAAGTATC 4410
Qy 8508 GAATATAGAAATCTTTAGCGTAGCCGGTGTGATGACACAGAAACAAACAAAGCATTTAC 8567
Db 4411 TTCAATTGAAAGTGAAGAAAGAAAAAATTTAGAAAGATACAGAAATTTAAAGAAAGGT 4470
Qy 8568 GAGATCAACAGATTAATCTTACACTGTAAATGAGAGAAACGATATCTCAACTTCGTC 8627
Db 4471 AGAATATATTAATTAAGTGTGATGCGCATTAAGGTTTGAAGAGAGATGATTAAGAA 4530
Qy 4471 AGAATATATTAATTAAGTGTGATGCGCATTAAGGTTTGAAGAGAGATGATTAAGAA 4530
Db 8628 AAAAGCTTTGCTAAAAATGAAAAATTAATGAAATGTAAGAAAGAACTGGAGAGGCTTACT 8687
Qy 4531 AGTATATGAT--TTTAAAGAGAGTATATTAAGCATGTTTAAAGAGAGATGGA--ATTAG 4587
Db 8688 CGAGCGGAAACAGCAGCCGTTGAAAAATTAACAAAGATCTACAGAGCATTTGGTTGC 8747
Qy 4588 GGATATGATATGAGAAAGTTTGAAGAGTATTAACAAACAAACCTTGAGAGAAAGTTGAATC 4647
Db 8748 AGGAAATTTGGAAATTTGAGATTAATTAAGAAAGATGCAAGAGTATATAGATTGTATG 8807
Qy 4648 CTTAAAGATTTTATCTAGTATGATGATGAGATGAGAAACAAATGAAGAAACAGAA 4707
Db 8808 AGTCAACGAGACGGAACCAAGAGAGCTTGTGCGAAAGATGATTTCTGTGAAAAA 8867
Qy 4708 AAAAGCTCAAGACTTAAGTTGGAAGAGATTTTAAAGAAAGAGTTTAAAGAAAGAAC 4767
Db 8868 TACAATTTACAGGAGAAACAAATATCCATTTGAAGATTAAGCCAGAAATTTGGAAACCG 8927
Qy 4768 AAAAGAAAAAATTAACAAAAAGAAAGTGAAGTTGATTTTAAAGATTAAGAAACCAAGAA 4827
Db 8928 AAGGTATTAATGATGCTTTGAATGAATGAACTGATGATGATCTACAGGAAAAAGTGTG 8987
Qy 4828 TGAATATGATGAAGTTGAAATGAAAGATGAAGATATGAAGAGATGTGAGAAAGATAT 4887
Db 8988 CTATGTGGAATTTGTATTTGAAAAATGTTGATGTAATATGATTTAAGAAAAATGTAGA 9047
Qy 4888 AAGAAAGATATTAAGAAAGATTAAGTTGGAAGATATGAATGAAGATATATGATGAAGATAT 4947
Db 9048 AGCCAAATTTGGAAGACATGCTATTTAGAAACCTACGAAACCAAGATATCAAGCATTT 9107
Qy 4948 AGGTGAAGACAAAGATGAAGTTATATGATTTATGTCAAAAAGAAAGAAACCATTTGAAA 5007
Db 9108 TACAAGAGCAAAAGTAAATATTTCTTGAAGAAAGAGACGCTCAAGTGAAGTGAATATC 9167
Qy 5008 GGTTAAGCGAAAAAGAAAAATTAAGAAAAAGTTGAAGAAAGTGTATGTGCTTAA 5067
Db 9168 GAATGTACACATTTCCATGAGATGATATTAATTAATTTGCAAGCAGTATGATCTTC 9227
Qy 5068 AAAACAGTGAACGAGATTAAGAAATATGTTTCAAAAAATTAATTAAGAAAGTTGATTAAG 5127
Db 9228 TCAATTAATTAACAAAAATTTCAAAAAATTAATTAATTTACTTACATCA 9274
Qy 5128 AGTATCTAAAGCTTTAGAAATCAAAAAATGATGTTACTATGTTTAA 5174

RESULT 10
US-08-232-463-14/c
/ Sequence 14, Application US/08232463

Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14

Query Match 0.9%; Score 87.6; DB 2; Length 7218;
Best Local Similarity 2.4%; Pred. No. 6.3e-08;
Matches 9; Conservative 251; Mismatches 120; Indels 0; Gaps 0;

QY 7098 GACTCCGTAAGCAAAATGTAACGCAAAAGCAAAATATTATTAAGCAAC 7157
DB 1445 GAAATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1386
QY 7158 AGCAGCAATGAGAGATGAGAGCAATCGGCAATGCTGTAATAATT 7217
DB 1385 RRR 1326
QY 7218 TGAACAGCTGTGAGATGAGAAAAATCTGAGAGAAAAAGCAAGTTTAAAC 7277
DB 1325 RRR 1266
QY 7278 TTAGCGAGTTAAACAAGACAGATGAGAAAAAGTAAATGATGTAAGAAAAATCTT 7337
DB 1265 RRR 1206
QY 7338 ACAATGAGAGATGATTTCTACAGAGATGATCTTGTAAAAAGCGATAGAGAGATCTA 7397
DB 1205 RRR 1146
QY 7398 GGAAGAGAGATTAAGCCATTGTGAAGACTTGTGATATTATGAGAAAAATGATAT 7457
DB 1145 RRR 1086

QY 7458 TACAACAGAGACAGAGATA 7477
DB 1085 RRRRRRRRRRRRRRRRRRR 1066

RESULT 11
US-08-956-171E-63
Sequence 63, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 8155 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-08-956-171E-63

Query Match 0.8%; Score 79.2; DB 3; Length 8155;
Best Local Similarity 41.3%; Pred. No. 3.4e-06;
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

QY 3782 TTCAAGTAAACAAGTCTTTGTTAAATAAGATGATTTGGAATGCAATTAATT 3841
DB 5651 TTAATTAATAAACCACTGACGTAATGAAATCAGCAATTTTAATAACAATTGCAAG 5710
QY 3842 AAGTAAGCTTGAAGTGAAGCACTCAAGTACAGAGAGAGAGTTGGAAGCAGT 3901
DB 5711 AATTCACAGTACGCAAGTGAAGATGAAGAAAAACAAGCAAGTATGCTGAAGCA 5770
QY 3902 AAGAAAGTGAAGACAGAAAGAAAGTTATCTATTGGGAACTTCTGCTTATCACTT 3961
DB 5771 ATACTGAATAATGTAAAGCAATCAAGCAATTCAGCAGCACTAATGCAAGAGT 5830
QY 3962 TGAACATGAAGTTCTGCAAAAATCAGAAAAATTAATCACTAGCAGAGAGATGAAAGCC 4021

Db	5831	ATGAAGCTAAAGCAAAATGCAAGACGAGCAATTAATGCGGTACACCAAAAGTTGTGAAGA	5890
Oy	4022	AAAAAATGATGTTGATGTCACCTGCTTACAAGCGACACCCAAAGTGCAGAGCTTTAA	4081
Db	5891	AACAAAGCGGCTAAAGATGAAATGTGATCAATTAACAAGCAACCGAACAATGTATTCAATA	5950
Oy	4082	ATTTCAGAGCTGGAAGATCAATGGAATCTGAGGGGCTACTGTGATCTGTGCCAAATTA	4141
Db	5951	ATGATCAAAACGCTCAACAAGAGAAAAAGAGCGCTATTCAACAATTACCAACAGCAG	6010
Oy	4142	ACAACAAGTAAATGCTCTATATTAGTGGAGATATACATAAGTTAATGAGCGGACG	4201
Db	6011	TTACGAGCGCGAAAAATATATTATACGTGCACACTGATGATTAATGTTAGATCAGCGGA	6070
Oy	4202	CAAAAGCTCTTTTAGCAACCACTCAAGTACTGCTGCAGTACGACGAGAGGACAAATTA	4261
Db	6071	AAGAGCGCTGGAAGAATTCATTTCAAAACAACCAACGACGACAGCGGTTAAATCAAAATG	6130
Oy	4262	GTTCTGAGCGGGATTAGGAAATTAACAAGGGGCTGTTCTGTCAATTAAGATGACAATG	4321
Db	6131	CTAAAAATGATGTTGATCAAGCTGTGACAACCTCAAAATCAAGCAATGATTAATCAACTG	6190
Oy	4322	ACGTGGAAGCTTAGCCTTGATTAATCTTCATCGAAGAGCTTAATGAATCAATGTCATTG	4381
Db	6191	GTCCTACAACTGMAAGAAAAATGAGAAAAAGATTGTGTTTAAAGCTTAAAGAAAAAG	6250
Oy	4382	CCAAAGATGTCMAAGAAAGTTCTGACTTAGCAAAAGATATCAGGCTTTACTAAATGGA	4441
Db	6251	CGTATCAAGATATCTTAAATGACAAACAATATGATGTTAAGCAATTTAAAGTCAAG	6310
Oy	4442	AAGATAAAAAATATTAGAAAGTCGTGTAATTAATGACATGGAATGTTATTATACGA	4501
Db	6311	CAGTTGCTGATATTCAAGGTATTACTGCAGATACAACTAAATTAAGATGTTG-----CGA	6364
Oy	4502	AGGAACAATCTGAAAAGCAAAAGAAAAAGAGAGCGGTCAATTGTAATCTGCTTTAT	4561
Db	6365	AAGATGAATTTAGCAACAAACGAAACGAAACGAAAGCGCTTATTGCAACAACTCAGATG	6424
Oy	4562	CGGTGCTGGAACGGATTAATCCGCTGAGAGTAGTAGTCAATTGCGAGTCAATCTGTAAAA	4621
Db	6425	CGACTCTGMAAGAAAAAGAACAGCAATCAACATGATGACGACACAAATTAACCAAGGTA	6484
Oy	4622	ATAAATTTAAAGCAGAAATGAGTAGACCAATTAAGGAAGCCGAGAGGATTAATTCATG	4681
Db	6485	ATCAAAATATTGAAAAATGCACAGTCAATCGATGATGTAAACATCGCAAAAGATATATGAA	6544
Oy	4682	CGAAACATGTAAATGTGAGAGCAAAATCATCTACTGTTGTTGTAATCGGCTTCTGAC	4741
Db	6545	TTCAAGCAATTGACCCCAATTCAAGCATCAACAGATGTTAAACCAATCAAGAGC-GGAA	6603
Oy	4742	TTGCTATCAGCAAGATGCTTTTTCAGGAATGGAATCTGAGCAATGGAACAATTTACAA	4801
Db	6604	TTGCTACTGAAATTCAAAATTAATTAATTAACCTGAATTAATTAATGAACT-----	6657
Oy	4802	ATGACACGATTCGAAAGGTGATAAAGAGAAATTTCTGCTGATTTCTTAAATGGAACG	4861
Db	6658	-----ACTTAATGAAGAAAAAGTAAAGTAACTTTGACACGATTAGACAGCATATTAAGAG	6712
Oy	4862	CAAAATATTCOAATCTTGCGGAGTAATGTTGCGGGAACCAATTCGCGGTTCTCTTTCTACG	4921
Db	6713	GTTTAAATTAATTAATATGACAGCACTAATAACAGGTGATTAATCTCTCTAAAGATACAG	6772
Oy	4922	CGGTGAGAGCTGCTTTTGCGAATTAATCTCTTCATTAATTAACCTCTGCTTTGATTACAG	4981
Db	6773	CAGTACAAAAAGTTCAACAACCTTCATGCAAAATCTGTTAAAGAAACCAACAGGTAAAAAG	6832
Oy	4982	GAAACAAGGTAAATCTTTTATGAGAAAGAAATTAACAAAGTCAATGTACAAGCTTTGATG	5041
Db	6833	AATTGATCAAGCTGACGTATTAAGAAAAACAATAATGAACCAACCAATGCAATC	6892
Oy	5042	ATTTCATATTACAAACGTTTCTGCTGAGGCGCTCAAGTATTAAAGACGCGTGAATCG	5101

D	b	6893	AACAAGAAATTAAATGATGCAAAACAAGAAGTTGATATCTGAATTAAATCAGCGAAAAACA	69552
O	y	5102	GAGGAATGATATCTGTCAATGCGGTTCTGATGAAACGGAAGCTTAGTTAGTATCTG	51615
D	b	6953	ATGTCGATCAATCATCAACAATGATATGTTGATTAATGCAAGTTAAAGAAAGAAAGCTA	70122
O	y	5162	AGTTGAAGAGTAAAGTTCTTCAATGTAATGCAAAAGATCAAAAACATTAATACAA	52212
D	b	7013	AAATTAATGCAAGTTAAAACATTTAGTGTGCAAAAAAGAGCTTAGCTTAATTAAG	70727
O	y	5222	TTGCGGAAATATGCAATGAGAAAGCGGCTGGAGTTGAGCAACGTTGCTCATCAAA	52815
D	b	7073	ATGCAATTAATGCTTAAGATTAACGAAGCGGATACTTAACGCATTCGACTTCAAGTAAA	71322
O	y	5282	ATATTGGAAAAACAATCAAGTTATAGCTATTGTAAAAAACAGTAAATTAACAACGCGAATG	53411
D	b	7133	TTGCTGAAGCGAAACAAAACTTGCTGATTTAAACAAACCGGGATCAAAATGTTATC	71922
O	y	5342	ATCAAGATAGAAAAAATATCAATGTACTGCAAAAGATTATATGACCAATATCTATAG	54015
D	b	7193	AAGCTACTTCTAAAGATGACATTAAGTTCAAAATTCATTAAGACTTATATTTAAACG	72522
O	y	5402	CAGTCGAGTTGGAGAGCAAAAGAGCCTGTGCAAGAGCTTCTGCAAGTACTACTCT	54615
D	b	7253	ATTAACACATTTCCAACGGTAAAAAAGAAATAGCTACAACGATTTATATGCTTATGAG	73122
O	y	5462	TGATTAAGACAGTTCTTCTCATGTTGATCAATCGATATTGACAAAGATTAGAGAG	55212
D	b	7313	ATCAGAAAGAAAAATATATTTCACTGACACTATATGCAACCAAGATGAAAAAGCAACAG	73727
O	y	5522	AAAAATATGGAATTAAGAAAAAGCAATGTTAATGTTTCTAGCTGAAATATCGAGTCAG	55815
D	b	7373	CAATTTAAGCAAGTTGACCAAAATG---TTCAACTGCAATTGAAAGCAATTAATATGCTG	74292
O	y	5582	TGTGTCACAAAATGCGACAGTGTCTTCCGAGAGCAATGCAACAGCTGCAAGTAGAGCTGAG	56411
D	b	7430	TGGATATATGGGACGGTTGATGATCAATTAACAAAGGTAAGACAGCAATGATGATCTATTC	74892
O	y	5642	TAGAGGTTAAATTAATTAACACAAATATCTTGCACATATTAATAATATAGTACTCAAAATG	57015
D	b	7490	AAGTATGATGCTACTGTTTAACCTTAAGCGAAACCAAGCTATTGAAGTTTAAGCAGAGATTA	75492
O	y	5702	TACGAAATGCTTTGGTAAAG-----CAAACTGATTCATCTA	57405
D	b	7550	CGAAAGATCATATTGATCAAAAGTGACACAGTTAACTGTGAAAGAAAAAACTGAACATTAG	76092
O	y	5741	TTAAAACAATTTGGAATTTGAGCTGGAAGTGGAGCTGGAAGCTGAGAGTCAGAGTTCTG	58005
D	b	7610	CAATGATTTAAACAATTTACAGATCACTTAACAAGGATATTAATGATGCAACCAACTG	76692
O	y	5801	TAGCAGTAATTAATATGTAATTAATACATAGCAGATTAATAATCATGCAAAAAATCACTG	58665
D	b	7670	CTGAAGTTGAAAAAGCGAAAGCTCAAGAGACTTGAAGCAATTTGATTAACATTCGAACT	77292
O	y	5861	CGAAGGGAATATGTCGAGTTATTAACAAGTCTGATGCGGTAAATGCTAATTAATGACAGAA	59205
D	b	7730	CAACAGAAAAACAAGATGCTATGCAAGATTAGAACTGCACTGACACAGATTGAAGAG	77892
O	y	5921	CAGTGTCTGAGAGTCCCGGTGCAACAATAGAGGCTCAACCAAGTGTGATGAATTAACG	59805
D	b	7790	GTTGTAATATGTCMAAGCTGATGCTCAACTGAAGAAAAAGAGCGTTTACGAAATCTTAG	78492
O	y	5981	GATCTACAAAAGCATATGTAABAAATTTCAACAGGATTTGCTAATGAAGAAAACAGATGAT	60405
D	b	7850	AAGACATTTTATCAAAAAGCAACTGAAGATATTTCTGATCAAACTACAAATGCAAAAATCG	79092
O	y	6041	ATATTACTACTCAAGGCAAGTATAGTAAAGGTGATTAAGTATTCAAAAATCTTATA	61005
D	b	7910	CTACTGTCAAAAATATAGTGCCTTGAACAACCTTAAGCAACAAGATTAATCTGAAAGTTA	79692
O	y	6101	TTAACGAAGCTTATCAAAAAAGAAAAATTAAGTAATAAAAAGAGATTGGTTACATA	61605
D	b	7970	AGAAAAATGCTTTGGAAGCATTCGAGAAAGTGTAAACCAAGCAAAATAGGAATATTTAAAA	80292

QY 6161 GTTCAGTACTACTACTT 6178
DB 8030 ATGCAGATGCAGATGCAT 8047

RESULT 12

US-08-781-986A-63
Sequence 63, Application US/08781986A
Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunach
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ. ID NO: 63:

SEQUENCE CHARACTERISTICS:

LENGTH: 8155 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-63

Query Match 0.8%; Score 79.2; DB 3; Length 8155;

Best Local Similarity 41.3%; Pred. No. 3.4e-06;

Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

QY 3782 TTCGAAGTAAACAAGTCTTGGTTGTTAAATAATGATATTTGAAATGCAATTAATTTA 3841
DB 5651 TTATATTAAGCACTGCGATGATGAAATTCACAGCAATTTTAAATTAACAATTGCAAG 5710
QY 3842 AAGTAATGCTTTGAGTGAAGAACTCAAGTAGCAGCAGAGCAGGTTTGAAGCAGTTA 3901
DB 5711 AGATTCAAGCTAGCCAGATGCAACAGATGAAGAAAAACAAGCAGCTGATGCGAAGCA 5770
QY 3902 AAGAAAGTGAAGCAAGAAAAAGTTATCTATTGGGAACCTTGTCTTATCAACTTAA 3961
DB 5771 ATACTGAAATGTTAAAGCAATCAAGCCATTTCAACAGCACTACTTAAAGCAAGTTG 5830
QY 3962 TGAACATGAAAGTTCTGCAAAAATAGAAAATATATACAGTAGCAGAGAACTGAAAGCC 4021
DB 5831 ATGAACTTAAGCAATTCAGAAAGCAGCAGTAAATGCGTAACCAAAAAGTTGTGAAGA 5890
QY 4022 AAAAAATGATGTTGATGTCCTGCTTATCAAGCGGACACCAAGTGACAGAGCTTTAA 4081
DB 5891 AACACGCGGCTAAGAAATGAATTTGATCAATTAACAACGCAAAACAATGTTATCAATA 5950
QY 4082 ATTTCAGAGCTGAAGAAATGAAATGAACTGTAGGGCTACTGTGACTGTTGCCAAATTA 4141

DB 5951 ATGATCAGAGGCTACACACAGAAAGAAAAGACGCTATTCAACATTAAGCAACAGAG 6010
QY 4142 ACAACAAGTAATGCTTCTATTAGTGGGAGATATACATGCTTAATGAGGGAGC 4201
DB 6011 TTACAGAGCGCAAAAATATATATACAGCTGCACATGATGAATATGTTAGTCAGGCA 6070
QY 4202 CAAGAAGCTTTTGAACAACCTCAAGTAGCTGCTGAGTGCAGCGGAGGAGCAATTA 4261
DB 6071 AAGAGCTGGAAGAAATTCATTCAGACGCAACGCAACGCAACGCGGTTAAATTAATG 6130
QY 4262 GTTCTGAGCGGGAATTGAAATTTATCAAGGGGCTGTTCTGTCATTAAGATTGA 4321
DB 6131 CTAATAAATGATGTTGATCAAGCTGACACTCAAAATCAAGCAATGATATACACTG 6190
QY 4322 ACGTGAAGCTAGCGTTGATTAATTTCCATCAGAAAGGCTAATGAATCAATGTCATTG 4381
DB 6191 GTGCTACCACTGAAGAGAAAAATGCAGCAAAAGATTGTTTAAAGCTTAAAGAAAAAG 6250
QY 4382 CCAAGATGTCAAAGGAAGTTCTGATCTAGCAAAAAGAAATTCAGGCTTTACTAAATGGA 4441
DB 6251 GGTATCAAGATATCTTAATATGCAACAACACTAATGATGTTACGCAAAATTAAGATCAG 6310
QY 4442 AAGATAAAAAATTTAGAAAGTCGTGTAATTAATACAGCTGGAATGTTATTAACGA 4501
DB 6311 CAGTTCGTGATATTCAAGGTATTACTGCAGTACACAAATTAAGATGTTG-----CGA 6364
QY 4502 AGGAACAATGAAAAAGCAAAAGAAAAAGAAAGAGCGGCTCATTTGAAATGCTGCTTAT 4561
DB 6365 AAGATGAATTGCAACAAAGCAAAAGCAAAAGAGCGCTTATTCACAAATGTCAGATG 6424
QY 4562 CGGTTCGTGGAAGGATTAATCCGCTGAGAGGTAGCTATTGCAATCAATCTGTTAAA 4621
DB 6425 CGACTCTGAAGAAAAAGAAAGCAAAATCAACAGTAGACGACAAATTAACACAAGTA 6484
QY 4622 ATAAATTTAAAGCAAGATTGAGTGAAGCAATTAAGAAAGCCGAGAGATTAATTAATCAG 4681
DB 6485 ATCAAAATATTGAATAATGACAGCTGATGATGATGATGAACATCAGCAAAAGATATGCA 6544
QY 4682 CGAAACATGTAATGAGAGCAAAATCATCTACTGTTGTTGTAATGCGCTTTCGAC 4741
DB 6545 TTCAAGCAATTAACCAATTCAAGCATCAACAGATGTTAAACGATGCAAGAGC-GGA 6603
QY 4742 TTGCTATCAGCAAAATGCTTTTTCAGGAATGGAATTCGACATGCGCAAGCTTATCA 4801
DB 6604 TTGCTATCAGAAATCAAAATTAATTAATCTGAATTAATTAATTAATGAGCT----- 6657
QY 4802 ATGACAGATTCGAAGGTGATTAAGAAAGAAATTTGCTGATTCCTTAAATGCAAGC 4861
DB 6658 -----ACTAATGAAGAAAAAGTTAACGATATTGACAGCTTAGACAGATATGAAGAG 6712
QY 4862 CAATTAATTCATTTCTGGGGTGAATGTTGCGGAACCATTCGCGGTTCTTTTCTACG 4921
DB 6713 GTTAAATTAATTAATGACGAACTACTACGAGTATTAATCTACTGCTTAAAGATACAG 6772
QY 4922 CGGTGAGAGCTCTTTTGGGAATTAATCTCTTCAATTAATTAATCTGTTGATTACAG 4981
DB 6773 CAGTCAAAAGATTCAACAACTTCAATGCAAAATCTGTTAAGAAACGAGAGTTAAAAAG 6832
QY 4982 GAACGAAGTAATCTTTTATGTAAGAAATATCAAAAGTCAATGATGACAGCTTTGAATG 5041
DB 6833 AATTGATTAAGCTGAGCTGATTAAGAAACAAATTAAGAAACAAACAAATGATCAGC 6892
QY 5042 ATTTCAATTAACAAGCTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGGCTGCAATG 5101
DB 6893 AACCAAGAAATTAATGATGCAAAACAAAGAGTTGATCTGAATTAATCAAGCGAAACAA 6952
QY 5102 GAGGAATGTAATCTGCAATCGGTTCGATGAAGCGAAGCTTTAGTTAGATGCTTG 5161
DB 6953 ATGTGATCAATCAATCAAAATGAATATGTTGATTAATCACTTAAGAAAGAAAGCTA 7012
QY 5162 AGTTGAAGAGTAAGTTCTTCAATGATGATCAAAAGATCAAAACAAATTAATTAACA 5221

Db 7013 AAATTAATGCGATTAAACATTTAGTGAATACAAAAAGATGCTTTAGCTAAATTTGAG 7072
Qy 5222 TTGCCGAAATGCAATGAGAGAAAAGCGCTGAGTTGAGCAACAGTTGCTCATCA 5281
Db 7073 ATGCATATATATCTTAAAGTAAAGGAGATACCTTACGCAATGCACTTCAAGTGA 7132
Qy 5282 ATATTGAAAAAATCAGTTATAGTTATTTGTAATAAAGTAAATTAACAAGGAAATG 5341
Db 7133 TTGCTGAAGCGAAACAAAACTTGCTGAATTAACAACTCGGATCAAAATGTTAATC 7192
Qy 5342 ATCAAGATGAAAAAATATCATGTGACCTGCAAGATTAATCTATGACCAATCTATAG 5401
Db 7193 AAGCTACTTAAAGATGACATTTGAAGTTCAATTCATTAATGACTTAATTAATTAACG 7252
Qy 5402 CAGTGAAGTTGAGAGAGAAAAGAGCGCTGTGCAAGAGAGCTTCTGCAAGTACTACCT 5461
Db 7253 ATTACACATTTCCACAGGTAAAGAAAAGATCAGCTACACAGATTTATATGCTTATGCG 7312
Qy 5462 TGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGAG 5521
Db 7313 ATCAGAGAAAAATATATATTTAGCTGACACTAATGCAACACAGATGAAAAAGCAAG 7372
Qy 5522 AAAATTAAGAAATTAAGAAAAAGCAAAATGTTAATGTTCTAGCTAAAAATACAGTCAAG 5581
Db 7373 CAATTAAGCAAGTTGACCAAAATG---TTCAAACTGCAATTAGAAAAGCAATTATATGCTG 7429
Qy 5582 TGGTCACAAAATGCGACAGTCTTCCGAGCAAGTGAAGCAAGCTGAGTGAAGAGCTGAG 5641
Db 7429 TGGATTAATGATACGTTGATGATGATGATTAACCAAGTTAAAGCAAGATTTAGCTTATTC 7489
Qy 5642 TAGCAGTTAATTAATTAACAAAAATCTTTCGACATTAATTAATTAATTAATTAATTA 5701
Db 7490 AAGTGAATGCTACTGTTAACTTAAGCGAACCAAGCTATTGAGTTAAAGCAGAAATG 7549
Qy 5702 TACGAAATGCTTTGTTAAAAAG-----CAATCTCTTATCTA 5740
Db 7550 CGAAAGAAATCTATTAATGATCAAAAGTGAACAGTTAATCTGTAAGAAAAAATGAAAGATTAAG 7609
Qy 5741 TTAATAACAATTTGAGAGTGAAGTGAAGTGAAGAGCTGAGAGAGCTGAGTGAAGTTCG 5800
Db 7610 CAATTAATTAACAAATTTACAGATCAAGCTAAACAAAGTATTAATCTATGACACACACAG 7669
Qy 5801 TAGCAGTAATTAAGTTGTAATTAATTAAGATTAAGCAATTAATTAATTAATTAATTA 5860
Db 7670 CTGAAGTTGAAAAAGCAAAAGCTCAAGAGCTGAAGCAATTTGATTAACAATTCAGACT 7729
Qy 5861 CGAAGGAAATGTCGAGATTTATTAACAGATCTGATGCGGTAAATTTCTAATTAATGACGAA 5920
Db 7730 CAACGAAAAAACAAGAAAGCTATCGAAGATTAAGAAATGCACTAGACCAAGATTTGAAGCG 7789
Qy 5921 CAGTGTGAGAGTCCGCGTGAAGCAATTAAGAGCTCAACAGTGAATGAATTAATTAACG 5980
Db 7790 GTGTAATATGTCACCGCTGATGCTACAACTGAAGAAAAAGAGCTTTACGAATGCTTTAG 7849
Qy 5981 GATCTACAAAAGCAATATGTAAGATTTCTACAGTATTTCTAAGAGAAAAGAGATTA 6040
Db 7850 AAGACATTTTATCAAAAAGCACTGAAGATATTTGATCAAACTCAATATGAGAAATCG 7909
Qy 6041 ATATTACTACTCAAGGCAAGTAAAGTAAAGTGAATTAAGTAAAGTAAAGTAAAGTAAAGT 6100
Db 7910 CTACTGTCAAAAATATGTCGCTTGAACAACTTAAGCAACAGTATTAATCTGAGTTA 7969
Qy 6101 TTAACGAGAGCTTATCAAAAAAGAAAAATTAAGTAAATTAAGTAAAGTAAAGTAAAGTAA 6160
Db 7970 AGAAAAATGCTTTGAGAGCAATCAAGAAATGCTTTAACAAGCAATTAAGAAATTAATTA 8029
Qy 6161 GTTACGACTACTACTT 6178
Db 8030 ATGCAAGATGCAATGAT 8047

; Sequence 1280, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIORITY FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIORITY FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1039)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

Query Match 0.8%; Score 76; DB 3; Length 1039;
Best Local Similarity 44.1%; Pred. No. 7.8e-06;
Matches 403; Conservative 0; Mismatches 507; Indels 4; Gaps 2;

Qy 144 AAAAAATGAAAAAAGATTAATGTTTATGACATTACTACAAACAGATTCAAGGAGAA 203
Db 107 AAAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 166
Qy 204 CGCTTTTACAGTTTATTAATGATTTGCTTTAAACGAAAAATTAATTAAGCAATTAATTT 263
Db 167 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 226
Qy 264 TGGGAAAAAGATGATCGGAGTAAATTAATCTTTTAACTTTGCAATGAAAAATTTGA 323
Db 227 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 286
Qy 324 AGTGAATGGAATTAACAGGAATTCGAGAAATTAATTAAGAGAAATTAATTAATTTCTT 383
Db 287 AAAAAATCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 346
Qy 384 AAGCTCGAAGAGATGCGAGTGAAGAAAAATGAGATTATCAATGCTGTTCTTTCAATTC 443
Db 347 AAAAAATTAATTAATTAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 406
Qy 444 TATTATTCAAAACAGATGATTTTAAGAGAGCTTTGAGAGAGCAACATGATTAAGT 503
Db 407 AAAAAAAGAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 466
Qy 504 TTTTAATGGAATCAATTCAGATGAGAAAAATTAATTAATTAATTAATTAATTAATTA 563
Db 467 CCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 526
Qy 564 CATTAAGGTAGAGAAAAATCAATGCTGTAAGGAGATGCGTTTATTAATGAGGAGATAT 623
Db 527 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 585
Qy 624 TAGATTTGAAGATCTCAATTAATTAAGAGAGATTAACAGATTTTAATAAATTTAGTCA 683
Db 586 AAAAAATCAATTTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 645
Qy 684 TATTAGATGGAATTAATTTCTGCTGACCGAGATTTAAAGCTTAACCAAGCAAAATC 743
Db 646 CTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 705
Qy 744 TGAAGATTAATTTCTTCAAGCTCAATGATTTCTCCCAAAAGC---TATGAGAAAAA 800
Db 706 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 765
Qy 801 TTCAACTGTGGAAGAGAAATAGAGAAATATGTAAGAAATTAACAAAGCAAAATATTGA 860

QY 7344 AGCAGTATTTCTACAGAGATACCTTGTATAAGCGATAGAGAGATACCTCAGGAGA 7403
DB 1438 AGAAGTAAAAAGAGAGTGTGCTACCACTTAAATAGAACTGTGGAACAGGCGAAGAGAA 1497
QY 7404 AGGAATTAAGCCATTGTGTAAGACTTCTGATATTAATGGAATAATGTGTA-----TAT 7457
DB 1498 GAGCAAGATCAAAATTAACGAAATATTGAAATTTAGAAAGAAATGCGATGAAAGTAA 1557
QY 7458 TACAACAGAGACAAGAAATATATCATCTCTACTGTGTGTTGGGAACGACGCTTGC 7517
DB 1558 TCAAAAAGTTGCGAGAAATTTAGAGAAATTAACGAACTGTATTTAATCTGATTTAGA 1617
QY 7518 TTCCGCATCAGAGACAGTGCAGTTACAAATTTAAAGAAATTCGAGTTACTGTTGA 7577
DB 1618 TAAAGTAGAGAAACGATGAATATTAGCGAGAAAGTTTAAAGAAACATGAATGATTA 1677
QY 7578 AAATCTTTTGTGAAGACGCTGAAAGTAATGTTAGATCGATATTACGGAATGT 7637
DB 1678 AGCATTATTTAGTGAATTTGATTAATGTAAGAAATTAACAAGAAATTTATTAACAG 1737
QY 7638 TGCTTT 7643
DB 1738 TATGTT 1743

RESULT 15
5231168-1
; Patent No. 5231168
; APPLICANT: DZIEGIEL, MORTEN, BORRE, MARTIN, JERSEN, SOREN,
; VUUST, JENS, RINECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALIE H.
; TITLE OF INVENTION: MALARIA ANTIGEN
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/409,658
; FILING DATE: 18-SEP-1989
; SEQ ID NO: 1:
; LENGTH: 3095
5231168-1

Query Match 0.7%; Score 64.6; DB 9; Length 3095;
Best Local Similarity 42.0%; Pred. No. 0.0023;
Matches 657; Conservative 0; Mismatches 884; Indels 22; Gaps 4;
QY 8163 AGTAA CAGGTTCTGTAATTGAGAGAGTTGAGATCAACCAAGCAGAGCTACTGTCAGG 8222
DB 849 AAAAAAGAAAATGTAATCTGAAGTTGTTGAGAAAAACAATTCACAAGATCAAGTTGA 908
QY 8223 TAAAACTATGTAGAAAGTTGAGAGAAATTTGTTCAAGACAAATCGATTGAATGCAAT 8282
DB 909 AGAAATTCAGATTAATGAGATGAATTTGAGATGTTCACTGAACAATTAATTTAGA 968
QY 8283 TTCTAAAGTAGAAGGTTGATGAAGATTAAGTAATCTCTGCTGATTCAGG 8342
DB 969 TATATTAACAGTTGATCCAGAAATAGTAGAAGTTGAAGAAATTCCTTCAGAACTACATGA 1028
QY 8343 AAATGAGAGAGAAATTCGCGAGCAGAGATGAATACTTCTACAGACAAGATGATATGTA 8402
DB 1029 AAATGAAGTGCTCATCCAGAAATTTGTAATTTGAGAAAGTTTTCCTGAACCAATCA 1088
QY 8403 ATCCGATGTTGTTTACGAAGCAAGATTAAGAAATTAATGATTAACAAAAATATAT 8462
DB 1089 AATATACGAATTTCAAGAAATTAATGAAGATTAAGTAAGTGAACATATTCACATGAAT 1148
QY 8463 TTCAGAGTCATGCTCTGCTTTAAATGATCAAAAGATGAAGGAATATAGATCTTT 8522
DB 1149 AGTAGAGTAGAAGAAATCTTCAGAGATGATTAAGAAATGAAGAAAGTTGAACATGAAT 1208
QY 8523 AGCGTAGCCGATGTGATGACAAGAGAAACAAGATTTAGAGATCAACAAGTT 8582
DB 1209 AGTAGAGTTGAAGAAATTTCTACAGAAATATTAAGT-----AAAAAGTCA 1256
QY 8583 AACTTCTACACTGTAAATGAGAGAAACGTATCTCAACTTCGTGCAAAAGCTTTGCTTA 8642

DB 1257 ACATGAATAATAGTAGAGTTGAAGAAATTTCTACAGAAAGATATATAAATGAAGAAAGTTGA 1316
QY 8643 AAATGAATAATATAGAAATGTAAAGAAATCTGAGAGACCTTAGTCGAGCGGAACAGC 8702
DB 1317 ACATGAATAATAGTAGAGTTGAAGAAATTTCTACAGAAAGATATAAATGAAGAAAGCTCAACA 1376
QY 8703 AGCGTTGAAAAATTATACAAAGAGTACTACAGAGCATTTGTTGCAAGAAATTTGGAAAT 8762
DB 1377 TGAATATGTAGAGTTGAAGAAATTTCTACAGAAAGATATAAATGAAGAAAGTTGAACATGA 1436
QY 8763 TGAAGATTAATTAAGAAACGATTCGACAGATTAATACGATTTGTAAGATCAACGAGACGG 8822
DB 1437 AATGTAGAAAGTTGAAGAAATTTCTACAGAAAGATATAAATGAAGAAAGCTCAACATGAAT 1496
QY 8823 AACCAAGAGAGTGTGCGAAGAAAGATGTAATTTCTGTGAAAAATACAAATTCAGGGA 8882
DB 1497 AGTAGAGTTGAAGAAATTTCTACAGAAAGATATAAATGAAGAAAGTTCAACATGAATGT 1556
QY 8883 AACCAATCATCATTTGAAGATTAAGCCAGAAATTTGGAAACCGAAAGTGTAAATGTAGA 8942
DB 1557 AGAAGTTGAAGAAATTTCTACAGAAAGATATAAATGAAGAAAGTCAACATGAATAGTGA 1616
QY 8943 TGCTTTGAATGAATCTTGAATGATCTACAGAGAAAGTGTGCTATGTGCAATTTGG 9002
DB 1617 GGTGAAGAAATTTCTACAGAAAGATATAAATGAAGAAAGTCAACATGAATAGTGA 1676
QY 9003 TATTGAAATGTGATGTAAATTAATGTGATTAAGAAATTTAGAAAGCCAAATTCGGAAG 9062
DB 1677 GATTGAAGAAATTTCTACAGAAAGATATAAATGAAGAAAGTCAACATGAATAGTGAAGT 1736
QY 9063 ACATGCTAATTTAGAAATCTAGAAACAAAGAAATATCAACATTTACAGAGCAAAAGT 9122
DB 1737 TGAAGAAATTTCTACAGAAAGATATAAATGAAGAAAGTCAACATGAATAGTGAAGTGA 1796
QY 9123 AAATATTTTGAAGAAAGAGCGTGCAGCTGCAGCTGCATATCGAATGTACATTTTC 9182
DB 1797 AGAAATTTCTCAGAAATTTGTTGAATTTGAAGAAAGTCAACAAATTAACAAATGA 1856
QY 9183 CAATGAGATGATATTAATAAATTTGCGAAAGCAGATGATCATCTCTCAATTAATTAACAA 9242
DB 1857 AAATATTTGAATCTATTAACCAAGAAAGAAAGAAATTAATTAATGTGTAAGAAAGAAAGC 1916
QY 9243 AAATTCAAAAATAATATATTAATTAATTAAGCATCAAGTGAATGCAATGTTCAATG 9302
DB 1917 AATTCACAAAGACCCGATGATCACTTAATTAAGAAATGAAGAAAGTATCTCCAAAGC 1976
QY 9303 GGTGCTGAACAGAGGTGCGAGAGCCAAAGCCAGCTTGTGTAAGAAATCAATTA 9362
DB 1977 ATCTGAAGTGAATCCACTTAACCAAGATATGTTCAATTAATTAAGTACAAAGAAATTA 2036
QY 9363 TAGAATCTAATATGTTGATTTAGCAGGAAATTAATAACAGAGGAAACATCAATGTATA 9422
DB 2037 ACCAAATTAAGAAAGCAACAGATGATGATGATGATGATTTGTAAGGATTT-----ATC 2096
QY 9423 TGCCGATATGATATAAATTAATATATAGTGAAGCAATTTCTAAGCTATTTGCGAGATGC 9482
DB 2097 AGAAGATGATATATGATGAAGAGATGATGATGATGATTTGTAAGGATTT-----ATC 2150
QY 9483 CAAAGATATGCTGAGCTGCTTGCGCAATGCGCACTTTGA-AAAAATGAAGTAAAT 9541
DB 2151 AAGAAAGATATGAAAGAAATTCATCAATTAATAAATTAAGAAAGTATCTTTTATATAC 2210
QY 9542 TTAATTAATGCAATCGAATTTTAATAAATTAATCTGCAAGATTTGGAAGGAAAGCTATA 9601
DB 2211 ATATATATCTCAAGAAATTTTAATAAAGTATCTCAACATTTGTAAGTGTATGATTA 2270
QY 9602 AAAAAACGTGCTAGAGATCTAA---TCAGGTAGATGTGATATACGATTAATATACATGC 9658
DB 2271 TGCAATATATGTGTATTAACAGTTGTGATGATTAAGAAATGCAAGAAAGATATATGT 2330
QY 9659 ATTCTTGAAGAAAGCATACAAAGAAATGACATATCAATCAAGAGAGAGAAAGGGA 9718
DB 2331 AATATTTTTCAAAAACATTTAATTAATTAACAAAAAATTAATTAATTAATTAATTAATTAATTA 2390

Qy 777 TCTCAAAAAAGCTATGGAAAAAATTCACGCTGGAGAAAGAAATGAAATATGTAAA 836
Db 39398 TCTATCATCTCTATTTATTAATAATCACTTATAGATTGATTAATTAATAATTAATAATTA 39339
Qy 837 AAGAAATACCAAGCAAAATATTGAAATCTG-ATGCTGTATTGGACAGATGAAATATTA 895
Db 39338 AATATATATTTATTTGATTAATTAATTAATTAATTTAAATTAATCAAAATTAATTAATTA 39279
Qy 896 AAATTAGTCGCAAAAGCTACAAATGGAGATTATTAAGAAAGAGGAAAAAGAACTT 955
Db 39278 AAATATAGAAATTTATTAATAATAATTAATTAATTAATTAATTAATTAATTAATTA 39219
Qy 956 ATACACCTCTTTAAGTTATCAATGTGAAAGCTCCGTAAAGTAATTAAGGAAAG 1015
Db 39218 ATTTAGCAAAATCTATTAATCACTATTAATTAAGAAAAAACCAAGAAATTAATAAG 39159
Qy 1016 TCATAGCAAAAGATTTGACATTACAGCTGAAGCAAAAGATTTCTATGAT 1065
Db 39158 AAGTAAATTAATGATGATGATTAATAAGTAATTAAGAAAGATTTTATTAAT 39109

RESULT 18

PCT-US93-07261-10
; Sequence 10, Application PC/TUS9307261
; GENERAL INFORMATION:
; TITLE OF INVENTION: P4EMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John H. C. Blaisdale
; STREET: One Gibraltar Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07940-1000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh 6.0.5
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07261
; FILING DATE: 19930805
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/927,531
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blaisdale, John H. C.
; REGISTRATION NUMBER: 31,895
; REFERENCE/DOCKET NUMBER: DX0288K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-822-7398
; TELEFAX: 201-822-7039
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4766 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; STRAIN: Malayan Camp
; IMMEDIATE SOURCE:
; CLONE: p2b1:p12-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..4766
; PCT-US93-07261-10

Query Match 0.6%; Score 62.2; DB 6; Length 4766;
Best Local Similarity 42.7%; Pred. No. 0.0083;
Matches 551; Conservative 0; Mismatches 728; Indels 12; Gaps 4;

Qy 8271 ATTGAATGCAATTTCTAAAGTAGAAGGTTGGATGAAGATTAAGTAATCTGCTAAATCTTC 8330
Db 2263 ATATTAAGAAACCTGATTTAGCAAAAGGTAAGAGAGTAATTAATAAGCAATGAAATCTT 2422
Qy 8331 TGTGTATCAGAAATGAGAGAAATTTCCGAGCAGAGATGAATCTTCTACAGCA 8390
Db 2423 AGAAGAAATATTAATGAACCTGATTTAGCAAAAGGTAAGAGATTAATTAAGACCGGA 2482
Qy 8391 AAGTAATCTGAATCCGATGCTGTTTACGAAGACAGATTAATGAATAATTAATGAATTAC 8450
Db 2483 AAATCTAGAAAGATTAATTAAGAACTGATTTAGCAAAAGGTAAGAGATTAATTAAGC 2542
Qy 8451 AAAAAATATATTTTCAGAGTCAATGCTCTTGCTTAATGATCAAAAGATGAAGCGAA 8510
Db 2543 ACGGAAATCTAAGAGATTAATTAAGAACTGATTTAGCAAAAGGTAAGAGATTAATTAAG 2602
Qy 8511 TATGATCTTTAGCGGTAGCCGGTGTGATGCAAG- - - -GACCAACAAAGCATTT 8565
Db 2603 TAAAGCAGCGAAATCTAGAAAGATTAATTAAGAACTGATTTAGCAAAAGGTAAGAG 2662
Qy 8566 ACGAGATCAAAAGATTAATCTTACAACTGTA- - -ATGAGAGAAACGATCTCAACTTCG 8624
Db 2663 AACTAATTAAGCAGCGAAATCTAGAAAGATTAATTAAGAACTGATTTAGCAAAAGGTA 2722
Qy 8625 TCGAAAGCTTTGGCTAAAAATGAAAAATATGAAATGTAAGAACTGAGAGAGCTT 8684
Db 2723 GGAAGTACTTAATTAAGCAGCGAAATCTAGAAAGATTAATTAAGAACTGATTTAGCAAA 2782
Qy 8685 AGTCGAGCGGAAACAGCAGCGGTTGAAATTAATTAAGAGATTAATTAAGAGATTAAG 8744
Db 2783 AGTAAAGAGTAACTAATTAAGCAGCGAAATCTAGAAAGATTAATTAAGAGATTA 2842
Qy 8745 TCGAGAAATTTGGAAATTTGAGATTAATTAAGAAACGATTCAGAGATTAATTAAGATTT 8804
Db 2843 TATGAATATTAATTAATTAATTAAGATTAATTAAGATTAATTAAGATTAATTAAG 2902
Qy 8805 AAGAGTCAACGAGACGAAACCAAGAGGCTTGTGCGAAAGATGATTTCTGTGAA 8864
Db 2903 AAAAAATTAAGATTAAGAAATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 2962
Qy 8865 AAATTAATTTAGAGGAAACAAATCATCTGATTAAGATTAAGCAGATTTGTTGAA 8924
Db 2963 AAATTAAGAAATTAAGAAATTAAGAGATTAAGAGATTAAGAAATTAAGAGATTA 3022
Qy 8925 CGAAGGTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3082
Db 3023 TAAAGATTTAGAAATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGATTA 3142
Qy 8985 TGCTATGCTGAAATTTGATTTGAAATGTTGATTAATTAATTAAGATTAAGATTAAG 9044
Db 3083 AGAATTAAGAAATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 3142
Qy 9045 AGAAGCAAAATCGAAGACATGCTATTTGAAATTAAGAGATTAAGAGATTAAGAGAT 9104
Db 3143 ATTTAGAAATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAT 3202
Qy 9105 ATTTAAGAGCAAAAGTAATTAATTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 9160
Db 3203 ACAAATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 3262
Qy 9161 CAATATGCAATGTCACATTTCCATGAGATG- - -GATTTAAATTTGCAAAAGCACTA 9218
Db 3263 AAATTAAGATCTGAAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAG 3322
Qy 9219 TGCAATCTTCAATTAATTAAGCAAAATTTCAAAATTAATTAATTAATTAATTAATTA 9278
Db 3323 TAAAGATCTGAAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTA 3382
Qy 9279 TGAATCAATGTAATGATTTGATGAGGTCGTAAGAGAGAGAGAGAGAGAGAGAGAGAG 9338
Db 3383 AGAATCTGAAGATTAAGAGATTAAGAGAGATTAAGAGATTAAGAGATTAAGAGATTA 3442
Qy 9339 AGTATGTAAGATCAAAATTAATTAAGATTAATTAATTAATTAATTAATTAATTAATTA 9398

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/719,822B
FILING DATE: 09/30/96
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17686US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3337 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
IMMEDIATE SOURCE:
CLONE: PMB3.3.1
US-08-719-822B-1

Query Match 0.6%; Score 61; DB 2; Length 3337;
Best Local Similarity 42.3%; Pred. No. 0.013;
Matches 456; Conservative 0; Mismatches 615; Indels 6; Gaps 2;

QY 5692 ACTCAAAATGTACGAAATGCTTGTGTAAGCAAAATCTCATCTTAATAACATT 5751
DB 504 ACAGAGAAGAGCACTACCTGCTGATATGAGAGAGAAAGAAAGCTGCCATATGAGACA 563
QY 5752 GGAATGGAGCTGGAGTTGGAGCTGAGAGAGCTGAGTACAGGTTCTGTAGAGTGAAT 5811
DB 564 GGGAGAGAGCTGGTATTTGGAGAGAGGAGAAAGCTGGCATTTGAGAGAGGAGAA 623
QY 5812 AAGATTGTAAATTAATCGATAGCAGAAATTAATCATGCAAAAAATCACTCGAAGGAAAT 5871
DB 624 GAACTGGGATTTGGAGAGAGGAGAAAGCTGCTGATTTGGAGAGAGGAGAAAGAACT 683
QY 5872 GTTCGAGTTATTAACAGAGTCTGATGCGGTAAATGCTTAATTAATGAGAGAACTGCTGCA 5931
DB 684 GGGGATTTGAGAGCAGAGGAGAAAGCTGATGATGAGACACAGAAAGATGTTCACTGAT 743
QY 5932 GTGGCCCGTCAGCAATAGAGGCTCAACCACTGTGA--TGAATTACAGATCTACA 5988
DB 744 GATGCCGAAACTGAAATGAGACACTGTATGTATGAGACACAGAAAGATGTTCACTGAT 803
QY 5989 AAGCATATGTAAAGATTCTACAGTATTTGCTAAAGAGAAACAGATGATTATATTACT 6048
DB 804 GGGGAGAGAAAGTACATGTTCTCTGCTCAAGAAAGTACACCTGCCATAGTATGAT 863
QY 6049 ACTCAAGGCAAGTAGATAAGTGTAGATAAGTAAATTCAAAAATCTTAATATTAACGA 6108
DB 864 GCCCTCTTTGGAAGTATTTGGATAAAGATATAATTTTGTATCAATTAAGATTTTCGAG 923
QY 6109 GACTTATCAAAAAAGAAAAATTAAGTAAATTAAGATTTTGTACCAATAGTTCACT 6168
DB 924 CCACTATTGCAAAATTTGCGGGGTACTGCTAAACATGTTACGGGACAAAGATTGCCA 983
QY 6169 ACTCATCTTTAAATCTTTATTTGGCAAAATGCGCTGTTTCAAGACAAAGCCGAG--TG 6225
DB 984 ATBAAACCTGTACATTACAGTGGAGAGAGACCCCGCAATACCAAGGAGAAAGATTA 1043
QY 6226 GCAGAGACTGTTAATATCAACAAAGTTTATGAGAGAAACAGAGCTCTTTGAGAAAAATCT 6285
DB 1044 GATGCCACTTCAGAGATGACTTCGCAATGTATGTTAAGAAATCTCCCGAGAAAGTAGAA 1103
QY 6286 ATATTAAATGCAAAACATTATTCTGTAAATTCAGAGATTAACAGAAATTCATCGGAGTA 6345

DB 1104 TTAGATTAAGATGAAAGAGCAACTGAGAGAAATCAACGAAATGGACCAACGAGAA 1163
QY 6346 GTAGTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6405
DB 1164 GAGCCAAACGAAAGATTGATGCTCCAGAGATGATTTTGCATTAGACGAAACTGC 1223
QY 6406 ATAAAGAAATTAACCAAGAGAGTTGTAAGAAATCAATATGTTGTAAGAGTTTCGA 6465
DB 1224 AGAAGAGAAACAGAGAAAGTATAGGAGAGAGAAAGAAAGAGAGAGAGAGAGAA 1283
QY 6466 GAAAGAGTGAATTAACAGCAATTTTAAGCAAGAAATTTCTTTTGAATCGAGTC 6525
DB 1284 GTATCAGAAAGAACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTTC 1343
QY 6526 GCAGAGCCGGGATAGAGAGCCGAGAGTGCAGAGACCTTTCCGTAAATCAATTTCCAGA 6585
DB 1344 GCATTAGATGGAATCTATTAGAAAGAAACGAAAGAACTGCAGAGAGAGAAACCTGA 1403
QY 6586 AAGACGAAAGTATGATGTAAGAGAGAAAGATTTTGTAAAGAAAGTGAATTACAGA 6645
DB 1404 GAGGAGAGAGAAACCTTAGAGAGAGAGAGAAACCTTAGAGAGAGAGAGAGAGAGAG 1463
QY 6646 AAACGTTATGTTCTGTTGCAATTTGAAATGCGCAGTCCGAGTGTGCAAAAGAGCT 6705
DB 1464 GAAAGAGATTAAGAGCAACTCCAGAGATGACTTCCAAATTAAGAGAAACCATCAGAGAA 1523
QY 6706 GGAATTTGAGAGCAGCTGCGAGTTTCAAAAGATGAATCAACAGAGAGAGAGAGTG 6762
DB 1524 GGAGAGAGGAGTG 1580

RESULT 23
US-09-458-1
Sequence 1, Application US/09092458
Patent No. 6231861
GENERAL INFORMATION:
APPLICANT: Barnwell, John
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,
TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby and Darby
STREET: 805 Third Ave.
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022-7513
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/092,458
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/719,821
FILING DATE: 09/30/96
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17686US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3337 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```

: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Plasmodium vivax
: IMMEDIATE SOURCE:
: CLONE: PMMB3.3.1
US-09-092-458-1

```

Query Match	0.6%	Score 61;	DB 3;	Length 3337;
Best Local Similarity	42.3%;	Pred. No. 0.013;		
Matches 456;	Conservative 0;	Mismatches 615;	Indels 6;	Gaps 2

OY	5692	ACTCAAAATGTATCGAAATGCTTTGGTAAAGCAAACTCTCATTCATTTAAACAAATT	5755
Db	504	ACAGAGAAAGGACTACTGTGTATATGGAGGAGAAAGAAAGCTGGCGATATGGAGCA	563
OY	5752	GGAAATTTGAGCTGAGATTGGAGCTGTAGAGAGCTGTAGAGTGCACAGTTCTGTAGCACTGAT	5811
Db	564	GGGGAAGAAAGCTGGATTTTGGAAACAGGGGAGAAAGAACTGGCGATTTTGGAAACAGGGAA	623
OY	5812	AAGATTGTAAATATATACGATAGCAGAAATTTAAATCATTCGCAAAAATCACTGGAGGGAAAT	5877
Db	624	GAAATCTGGCGATTGGAAACAGGGGAGAAAGCTGTGTATTGGAGACGGGGAGAAACT	683
OY	5872	GTGGAGATTATTAACAGAGTCTGTATCCGGTAATTGGCTAATTATGACAGAAACAGTGTCTGGA	5933
Db	684	GGCATTTTGGAAAGCGGGGAGAAACTGGAGATTCGGGAACTGGAAAGAGAGCACTGGA	743
OY	5932	GTGGCCCGTCAGCAATAGAGCTTCAA CCACTGTGA--TGAAATTAACAGATCTCA	5988
Db	744	GATGGGAAACTGAAATAGAGCAACTGTGTATGTAGACACAGGAAGATAGTTCACTGAT	803
OY	5989	AAAGCATATGTAAAGATTCTACAGTGAATGCTAAAGAGAAACAGATGATATATTACT	6046
Db	804	GGAGCAGAAAGTACATGTTCTCTGTCTCAAGAAATGTACAACTGGCCGATATGTATGAT	863
OY	6049	ACTCAAGGCAAGTAGATTAAGTGTAGATTAAGTATTCAAAAATCTTAATTTAAACGA	6108
Db	864	GCCCTCTTTGGAGATATTGTGGATTAAGATATATTTTGTGATCATATTAAAGATTCCAG	923
OY	6109	GACTTATTCACAAAAAGAAAAATAGTATTAATAAAAAAGATTGTTACCAATATGTTCACT	6166
Db	924	CCACTATTGCAACAAATTTGTGGCGGGTCTGTCTAAACATGTTACCGGACAAAGATTGGCA	983
OY	6169	ACTCACTTTTAAATCTTTATTTGGCAATGCCGTGTTCAGACAGACCGGAG---TG	6222
Db	984	ATGAAGACTGTACCTTACGATGTGCAGAGAGAGCCCGCAAGTACACAGCGGAAGATTA	1046
OY	6226	GCAGGAATGTTAATATCAACAAGTTTATGGAGAAACAGAAAGCTCTTGTAGAAATTTCT	6285
Db	1044	GATGCCACTCCAGAGAGTAGACTCCCATTTAGATGTTTACAGATCTCCGAGGAAGTAGAA	1105
OY	6286	ATATTAATATGCAAAACATTATTTCTGTAAATATAGAGATTTACAGAAATTTCAATCGAGTA	6344
Db	1104	TTAGTATTAGATGAAGAGGCACTBAAGAGAGATCA CGBAAGTGGACCAACGGAAGA	1167
OY	6346	GTAGGTTCTGTGGTGTGTGGAAATGTAGAGAGTAGAGCTTCTTCGATACCAATATT	6405
Db	1164	GGACCAACCGAAGATTTGATGCCACTCCAGAGATGTGATTTCCGATTAGACGAAACGTC	1222
OY	6406	ATAAAAAGAAATACCAAGACAAAGATTGAAAACTACATGTCTGATGAAGTTTCGA	6466
Db	1224	AGAAAGAGAAACAGAGAAACGTAAAGGAGAGAGAAACAGAGAAAGCTGCAGAGAGAGA	1283
OY	6466	GAAAGAGCTGAAATTAACAGAGATTCTTAAGCAAGGAATTTCTCTTTTGGAGTCGAGTC	6522
Db	1284	GTAATCAGAGAAACCTCCAGAGAGAGAAAGATGTTAGAGCACTCCAGAGATGATATTC	1344
OY	6526	GCAGCAGCCGGGATAGAGACCGAGATGGCAGGAACCGTTTCCGTAAATCAATTTGACAGA	6585
Db	1344	GCATTAGATGGAATCTACATTAGAGAAACCGAGAGAACTGCAGAGAGAGAGAAACCTGA	1400

QY 6586 AAGACGGAAGTAGATGTCGAAAGCAAAAGTTTGGTAAAAAAGCTGAGATTTACGCA 66145

Db 1404 GAGGGAGAGAAACCGTAGAGGAGAAAMAAACCTAGAGGAGAAAGAACTGCAAGAGGA 14638

QY 6646 AAACGTTATAGTCTCTGTTGCCATTGGAAATGCGCAGCTCGAGTGGCTGCCAAAAGAGCT 67058

Db 1464 GAAAGAGGTTAGAGGCCAATCTCCAGAGAGATATCTCCATTATGAAGAACCTTCACGAGAA 15238

QY 6706 GGAATTGGACGACGAGTGGCAGTTACCAAGATGATTCMAACACGAGAGCAAGAGTG 6762

Db 1524 GGAAGAGGGGAGAGGAAGGAAGAGGGGAAAGGAAGAGAGAGCGTTAGTAGCGAGTG 1580

RESULT 24
US-08-719-821C-1

GENERAL INFORMATION:
APPLICANT: Barnwell, John
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,
TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby and Darby
STREET: 805 Third Ave.
CITY: New York
STATE: New York
COUNTRY: USA

DATE: 1002-7513
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/719,821C
FILING DATE: 09/30/96
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3337 base pairs

TYPE: nucleic acid

STANDARDNESS, double

STANDARDS: 14-30000

TOPOLOGY: linear

MOLECULE TYPE: I

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

OPCANTSM. PJ 94

IMEDIATE CORRECT:

US-08-719-821C-1

Query Match	0.6%	Score 61	DB 3	Length 3357
Best Local Similarity	42.3%	Pred. No. 0.013		
Matches 456; Conservative	0	Mismatches 615	Indels 6	Gaps 2

Oy	5692	ACTGAAATATGACGAATGCTTTGGTAAAAAGCAAACTCATTCATCTATTAAAAACATT	5751
Db	504	ACAGAGAGAGACTTACTCGTGATATGAGAGAGAGAGAGAGCTGGCGATATATGAAACA	563
Oy	5752	GGAATTGAGCTGAGATTGAGAGCTGAGAGAGCTGAGATGACAGTTCTGTAGCAGTAAT	5811
Db	564	GGGAGAGAGCTGTGTGATTGGAGACAGGGGAAATACTGGCGATTTTGGAAAGCAGGGGAA	623

```

QY 5812 AAGATTGTAATATACATGACAGATTAAATCATGCAAAAATACCTGCCAGAGGAAT 5871
DB 624 GAATCTGGCGATTGGAAGCAGGGGAAAGCTGCTGATTGGAGCAGGGGAAAGAACT 683
QY 5872 GTCCGAGTTATACAGAGCTGATGCGGTAATGCTAATTTATGAGGAACAGTGTCTGA 5931
DB 684 GCGGATTTGGAAGCAGGGGAAAGAACTGAGATGCGGAACTGAAGAGAGCAACTGGA 743
QY 5932 GTGCCCCGTGACAGATAGAGACCTCAACCACTGTGAA--TGAATTAACAGATCTCA 5988
DB 744 GATGCGGAAATCGAAATATGAGCAACTGTGATGTAGACAGAAAGATAGTCACTGAT 803
QY 5989 AAGCATATGTAATAAGATTCTACAGTATGCTTAAAGAAAGAAACAGATATATTTACT 6048
DB 804 GGAGCAGAAAAAGTCAATGTTCTGCTCAAGAAAAATGTCAACCTGCCATATGATGAT 863
QY 6049 ACTCAAGGCGCACTAGATTAAGTGTAGTAAAGTATTAATAAATCTTAATTTAAGCA 6108
DB 864 GCCCTCTTTGGAAGTATTTTGGATTAAGATTAATTTTGAATCATTTAAAGATTTCCAG 923
QY 6109 GACTTATCAAAAAAAGAAAAATTAAGTATTAATAAAGATTTGTTACCAATAGTTCAGCT 6168
DB 924 CCACTATTCGAACAAATTTGCGCGGTACTGCTAAACATTTACGGGACAGAAATTCGA 983
QY 6169 ACTCATCTTTAAATCTTTATTTGGCAATGCGCGCTGTTCAAGACAAAGCCGAG--TG 6225
DB 984 ATGAACCTGTACCAATTACAGTGGCAGAAAGCCCGCAAGTACCAACGGAAAGATTA 1043
QY 6226 GCAAGAACCTGTTAATTCACAAAGTTTATGAGAAACAGAACTCTGTAGAAAAATTTCT 6285
DB 1044 GATGCACTCCAGAGAGTACTTGCATTGATGTTACAGAAATCTCCAGAGAAAGTAA 1103
QY 6286 ATATTAAATGCAAAACATTATTCTGTAATAATCAGAGATTAACAGAAATCAATCGAGTA 6345
DB 1104 TTAGTTTATGATGAAAGCACTGAAAGAAATCAACGAAAGTGGACCAACGAAAGAA 1163
QY 6346 GTAGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6405
DB 1164 GGACCAACCGAAGAAATTAGATGCACTCCAGAGAGTGAATTTGCAATTAGACGAACTGC 1223
QY 6406 ATTTAAAGAAATTCACAGACAAAGTTGAAAAATCTCAATGCTGTATGAAGTTTCCGA 6465
DB 1224 AGAAGAGAAACAGAAAGAACTGAGGAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1283
QY 6466 GAAAGAGCTGAATTTACAGCAATCTTAAGCAAGAAATTTCCCTTTTGAAGTGGAGTC 6525
DB 1284 GTATTCGAAGAACTCCAGAGAGAAAGAGAGTTAGAGCACTCCAGAGATGATTTTC 1343
QY 6526 GCAGCAGCCGGGTAGAGCCGAGTGGCAAGAACGTTTCCGTAATCAATTTGACAGA 6585
DB 1344 GCATTTGATGGAACCTAATTAGAAAGAACCGAAGAAATGACAGAAAGAAAGAAAGCCTA 1403
QY 6586 AAGACGGAAGTATGATGGAAGAAAGAAAGATTTTGTAAAAAAAGCTGATTAACGA 6645
DB 1404 GAGGGAAGAAAGAAACCTGAGGAGAAAGAAACGTAAGGAGAAAGAAAGCTGACAGAGA 1463
QY 6646 AAACGTTATAGTTCTGTGCAATTTGAAATGCCGAGTGGAGTGGCTGCAAAAGAGCT 6705
DB 1464 GAAAGAGATTTAGAGCACTCCAGAGATGATCTTCAATTTAGAAAGCAATCCAGAGAA 1523
QY 6706 GGAATTTGAGCAGCAGTGGCAGTTTACCAAGATGAATCAAAACAGAGAGCAAGATG 6762
DB 1524 GGAAGAGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1580

```

RESULT 25
 US-09-754-947-2
 ; Sequence 2, Application US/09754947
 ; Patent No. 6828110
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, Bruce Andrew
 ; APPLICANT: Flores, Becky Mar

```

; APPLICANT: ValKire, Gunars Edvin
; APPLICANT: Biocite Diagnostics, Inc.
; TITLE OF INVENTION: Assays for Detection of Bacillus Anthracis
; FILE REFERENCE: 014907-003310US
; CURRENT APPLICATION NUMBER: US/09/754,947
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,901
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.1
; LENGTH: 2370
; TYPE: DNA
; ORGANISM: Bacillus anthracis
; FEATURE:
; OTHER INFORMATION: surface array protein (SAP)
US-09-754-947-2

Query Match 0.6%; Score 60.2; DB 3; Length 2370;
Best Local Similarity 41.2%; Pred. No. 0.017; Indels 3; Gaps 1;
Matches 501; Conservative 0; Mismatches 713;

QY 5352 AAAAAATATCAATGTCAGTCAAAAGATTAATATGACCAATATCTATAGCAGTGGAGT 5411
DB 969 AAAAGTACTTCACTACTGTAAGAGCTTTATATAAAAAGACGCTAAAGTAGTACTGA 1028
QY 5412 TGGAGACAAAGAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTTACCTGAATAGAC 5471
DB 1029 AAGTAAAGAAAGAAAGTTCGTGCAAGAGTGTGCAAGTACTTCAATCTTCACTGAGC 1088
QY 5472 AGTTCTTCTCATATGTCATCAAACTGATATTGACAAAGATTAGAGAAAGAAATATATG 5531
DB 1089 AGTTCAGAAACAAATTAAGCTGACTTACTTCTTAAAGTTTCAACAAACAAATTAAGT 1148
QY 5532 AAATAGGAAAGGCAATGTTAATGTTCTAGCTGAAATATGAGTCAAGTGTGACAAA 5591
DB 1149 ---TTACGAAGCGCAACAGCTTACGTTCAAGTAAATGAAAGATCAATTTAAGCAGT 1205
QY 5592 TGGCAGTGTCTTCCGAGACAGTGAACAGCTGCAAGTGAAGCTGAGAGTGAAGTAA 5651
DB 1206 AACCACTGAAAGTGAATATGATGCTTTAAACACAGAAAGTGTGATGATTAAGC 1265
QY 5652 TAAATTCACAAATTAATCTTGCACATTAATAAATAGTACTCAAAATGTACGAAATGC 5711
DB 1266 TACTGTAAAGTAACTGATTAATTCGACAGAAAGACCAAGTAAAGTAACTGTAAAGA 1325
QY 5712 TTTGGTAAAGCAATCTCATTCATCTATTAAACAAATGGAAATTTGAGCTGAGTGG 5771
DB 1326 TTCAAAAGGTAAAGCACTTGTTCACACAGTTGAATTTGAAGCTTGTGCTCAAAAGC 1385
QY 5772 AGCTGAGAGCTGAGTGAACAGTGTGTGTAGCAAGTAAATGAATTTAATTAATGAT 5831
DB 1386 AATGAAGACATTAATTTAGAAAAAACTAACTAGCGCTTTCTTACAAAAGATGTACAGA 1445
QY 5832 AGCAATTAATTAATTCATGCAAAATGACGCGAAGGAAATGTCCGAGTTATTACAGATC 5891
DB 1446 TTTAAAGTAAAGCTCCAGTCTAGATCAATACGTTAAAGTTTACAGCTCTGTAAAC 1505
QY 5892 TGATGCGGTAAATTTGTAATTTATGAGAAACAGTGTGAGTGGCCCGTGCACATAGG 5951
DB 1506 AGTGAAGTAACTGTGTAAGAGTGTAAAGATTTAAAGAACAAATTTAGAAAGCTAAATA 1565
QY 5952 AGCTTCAACAGTGTGAATGAATTAACAGATCTACAAAGCATATGTAAAAAGATTCTAC 6011
DB 1566 TGTGAACAGAGATTAGTTCTGAATGACAGACAGTCAAGAAAGCTGTGAATTAATACAGTTGT 1625
QY 6012 AGTGAATGCTAAAGAAAGAAACAGATTAATTTACTACTCAAGGCGCAAGTATGATTAAGT 6071
DB 1626 ATTTAATGCAAAATCTGTGTAAGAAAGAAAGCAAAAGCTTAATTTAGCTTGTGAATTAAGC 1685
QY 6072 GGTAGATAAGTATTCAAAATCTTAATTTAATTAACGAAAGCTTAATCAAAAAAGAAAT 6131
DB 1686 TCCAGTGCATCTCTAAATTTTGAAGTTGTGTGTTTGAACAGATTTGATTAATATGTG 1745

```


QY 519 TCCAGTAGAGAAAAGTAAATTCATGATCCGAATGGAAGCAATTACGGTAGAAG 578
DB 1452 TTTTAAATTCAAAAGTGAAGAAAAGATPAACACACCAATTCAACATTAAGAAAGAAAATCA 1511
QY 579 AAAAATCATGCTGTGAAGGATCGGTTTATATGCGCGGATATTAGATTGAAGATAC 638
DB 1512 AGTAATGTTGTCAGAGAAAATAATTCAGATTATTCAAGAGATTAATTAATAAAATTAAG 1571
QY 639 TGCATTAATAAGACAGGAATTAAGATTTTAAATTTAGTCAATATTAGTCAAT 698
DB 1572 CAAAAGATTAACATGAAATGTTAGATTAATTAAGAAATTAACAAATTAATGATATAA 1631
QY 699 AATTTGCTGTCAGCCGAGATTTAAAGCTCCAGACAAATTCGAGATTTATTTCT 758
DB 1632 AAATGTTGATGATTAATAAAATGTTGAGATTTAAAGTGTGAGATTAATAAAAGT 1691
QY 759 TTCAGCTCATAGATTCCTCAAAAAGCTATGGAAAAATTCATCTGTTGAAAAG 818
DB 1632 TGATATATTAACAATGTTGATGCTATTAATAATGTTGATGTTATTAATAATGTTGATG 1751
QY 819 AATAGAGATATGTAAGAAGAAATACCAAGCAATATTGAATGCTGATTTGA 878
DB 1752 TATTAATAATGTTGATGCTATTAACAATGTTGAGATTAATAACAATGCTGAGATACAA 1811
QY 879 AGCAGATGAATTAATAATTAATGTCGAAAGCTCAAAATGGAGATTTATTAAGAA 938
DB 1812 TAATGCTGAGATTAATAACAATGTTGCGATTAATAACAATCTGATGATTAATACAACT 1871
QY 939 AGGGGAAAAGA 950
DB 1872 TGAACATATAGA 1883

RESULT 28
US-09-806-708B-22
Sequence 22, Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(1141)
OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. PAEI promoters
US-09-806-708B-22

Query Match 0.6%; Score 59.4; DB 3; Length 1141;
Best Local Similarity 10.4%; Pred. No. 0.019;
Matches 80; Conservative 287; Mismatches 405; Indels 0; Gaps 0;
QY 858 TGAATCTGATCTGTATTGAGACAGATGGAATATATAAATTAAGTCGAAAGCTACAA 917
DB 94 TSBRYABARKKDKITVAVMTMTKMKGTGRHRYMRYABMDVDDHNYLAAMNAWTTC 153
QY 918 TGGGAGATTATTAAGAAAGAGGAAAAAAGAACTTATACATCTCTTAAGTTATC 977
DB 154 MKDKDKRTTMMWKNNATGWDMDTKYHMMNNNGVYVWVRKYKTRDRMSBRNNGM 213
QY 978 AGATGTGAAGCTCCGTAGAGTAATAAGAAAGAAAGTATGGAAGAAAGTTCGAT 1037
DB 214 BMMKMSYDVTTYTWVWDDMKCKKVRWVTRGRMRWVAVMTAARRRNNNGTBMAY 273

QY 1038 TACAGCTGAAGCAAGAATTCTATGATGCAACTTAGTACTTAAGCTTGCAAAAGCACTC 1097
DB 274 RRTMNNNNNNNAKMKRAKTMGRABVNSTCTTMSKTTKATSCANNCRBDANKD 333
QY 1098 TTTTACGTTTGTACAGGTTCTATTCTCTTAATTAATTAATGATTTAGGTTATT 1157
DB 334 HKMKMSAAMGVNN 393
QY 1158 GACAAGTAACTCCAGTGTCTGTTATGAAAAAGATCCCAAGATCCAGACAAAGAA 1217
DB 394 RKTWNNNNNGTMMKRMWAMTWKMDMBGTYNNNNNGRYYGWTCKKKMTYYKWTAN 453
QY 1218 GCGAAATTCATCTTACAGTGGAGTAAGCAACTAGGAGAGCACTTCTCCATT 1277
DB 454 NCKRPMWHTKCTHNNTTMMKTKYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 513
QY 1278 AAAAATTACCAATTAATTAATTTGAGAAAGCAAGTGAAGTCAATCGAGCGG 1337
DB 514 NNNMMDYMKACTMYKYBVCCKMNNVYAAWYTKSMYTSRYWKTNNSWRBDTRSM 573
QY 1338 ATATATTTCTCAAAAAGTAATCCATATGTAATTAAGAGAGAAAGTAAATCGAAGG 1397
DB 574 GRANNYABABHYGYKMTTRMBWSHTWBHBAAGAHYMBWYBAKCHCMRAWYAKYA 633
QY 1398 AAGACAGATATTAATCTCAAAATCGAAATGAAATGATGCTGTTCTGTTGGAAC 1457
DB 634 GAGGNN 693
QY 1458 GATGAGATTCATTAAGTAGCTCTTCAATGTTGTCGAGAGAGAGAAATTAATC 1517
DB 694 WGCNNNAATDTRRTTMMKNN 753
QY 1518 TTCGTCAGATGCTTAAGAGCAAAAGTGAATCAGAAACGATGATTAATGTCAG 1577
DB 754 KWAMGMAADAAABTTDKRNGAYTKYTTNNNNNTYRGVNTTAARDMANNNNNNNNNN 813
QY 1578 AAGTAAACGATTAATTCATTCGAGCTGCTGTAAGAAGTGGATGGGGGAT 1629
DB 814 NNNNGMSDMWVWYVANYGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 865

RESULT 29
US-08-682-517-7
Sequence 7, Application US/08682517
Patent No. 5874267
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Expression of surface layer proteins
NUMBER OF SEQUENCES: 25
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/682,517
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-682-517-7

Query Match 0.6%; Score 59.4; DB 2; Length 4197;
Best Local Similarity 42.0%; Pred. No. 0.03;
Matches 826; Conservative 0; Mismatches 1111; Indels 30; Gaps 7;
QY 5830 ATAGCAATTAATAATCATGCAAAAATCACTGGAAGGAAATGTCGAGATTATACAGAG 5889

Db 1916 AAAGCAAAACAATCTCTGTAAAAACAATGATTAAGGAAAGCTACATTTGTTTCGGT 1975
 Qy 5890 TCTGATCGGTATTTGCTAATTTATGACAGAAACAGTGTCTGAGTGCCCGTGCACAAATA 5949
 Db 1976 TCTGATACGTAAGATTAATGCAACACCAATTTGCTGATGATTAATTAATCTTCTGAT 2035
 Qy 5950 GGAGCCTCAACCAAGTGTGAATAAATTACAGATCTCAAAAAGCATATGTAAGATTTCT 6009
 Db 2036 GCAAAAACAAGCGACCTTGATGAAGTGAAACAAAAGAGTTGACCAATCTCTTACTTC 2095
 Qy 6010 ACAGATGTTGCTAAAGAAAGAAACAGATGATTAATTAATTAATCAAGGAGATTAATAA 6069
 Db 2096 CAAAGACCATATCTTGATGAGCTCAGCTATCAAGCATACAAAATTCAGATCTTAATAA 2155
 Qy 6070 GTGTGATTAAGATTTCAAAAATCTTAATATTAAGAAAGCTTATCAAAAAGAAAA 6129
 Db 2156 GCTGTACCTAATTTGATGAGTTCTGAAACTGCAGATTTTGGCAGCAAAATTAATAACCA 2215
 Qy 6130 ATAAATTAATTAAGAAAGATTTGTTACCAATAGTTCACTCATTAATTTAAATCTTTA 6189
 Db 2216 AGCGCAAAAAGTAAGTCTGTAATCTTAATAAGAAAGCACTTAATACATCTACAAATCT 2275
 Qy 6190 TTGGCAATGCGCGTGTGTTACAGACAGCCGAGTGGCAGAGACTGTTAATATCAACAG 6249
 Db 2276 GGTGCTAAATGATTAATTAAGTATGATACCAAGTTATCTCAAAAATCTGATGACAGTA 2335
 Qy 6250 GTTTATGGAAGAAACAGAAAGCTCTGTGAGAAATTTCTAATTAATGCAAAAACATTAATCT 6309
 Db 2336 ACTTATGAAGCTATCTTATCTTCTACAGAACTGTTATTAACCTGCTTAAGAAATTAAGA 2395
 Qy 6310 GTAAATCAAGAGATTAACAGAAATTCATCGAGTATGAGTTCTGTTGGTGTGTGGA 6369
 Db 2396 GTTACTTCAGTGGATGATTAACCACTGCTGTTAAAGTAAATGTTACAGGTATTTCTGT 2455
 Qy 6370 AATGTAAGATGAGAGCTTCTCTGTATACCAATATTAATAAGAAATTCACAAAGCAAGA 6429
 Db 2456 AATACAGACGTAAAGACATATGATTAATGCTTAAGAAAGCTACAGCTACATTCACACT 2515
 Qy 6430 GTTGGAAAACTACATGCTGATGAAAGTTTCGAGAAAGAGCTGAATTAATCACAGAT 6489
 Db 2516 ACAATTAAGATTCAAAATCTTTACATGCTGATGATGATCACTAA---TTCAATACACTGAT 2572
 Qy 6490 TCTAAGCAAGAAATTTCTCTTTTGGAGTCCGAGTCCGACAGCCGGGTAGAGACCGGA 6549
 Db 2573 TCTGATTCAAAACAGAACTCTATTTGGTGTGCTGTAAGAAACCCAGTAAATATCTGCT 2632
 Qy 6550 GTGGCAGAAACCGTTTCCGTAATCAATTTGCAAGAAAGACGAAATGATGTGAAGAA 6609
 Db 2633 GTATCAGGCAAAACATTAATTAATTCGAGGCTAAATGTAATGAAATTTTGGTAAACG 2692
 Qy 6610 GCA---AAGATTTTGGTAAAAAAGCTGAGATTAACAGCAAAAAGCTTAATGTTCTGTTGC 6665
 Db 2693 GATGAGGAAGCATTTTAATCAATATGCAATGCACTGAAGGCCAAAAGTAACAATTCATAT 2752
 Qy 6666 AA--TTGAAATGCGCAGTCCGAGTCCGAGTCCGAAAGAGCTGAAATTTGAGCAGCAGTG 6723
 Db 2753 AATGTAAGATGATGATACAGTACATTTAAAGTAATTAAGTGAATTTATTTCTTCAACTGA 2812
 Qy 6724 GAGATTAACCAAGATGATCAACACAGAGCAAGAGTGAATAATTC---TAAATTA 6778
 Db 2813 GCTATCAAAACAGTTGCTCCAAACAACAGAGCTCCCAATCACTGCGCATTAACATTA 2872
 Qy 6779 TGACTGAAACAAGT-----TAGATGTAATAGCAAGAAATGAGATTAATCAAGTACT 6831
 Db 2873 ACAACAGCAAGCTGATGATTTAGTTGATTTTAACAATCTGCACTAACCTTAGAATTTCA 2932
 Qy 6832 GGAATCGGTTACGCGAGCTGGAATTTCTTGACGCGAGATTAATCTGAGTGTGTTCTGTC 6891
 Db 2933 TTAGCTGAATGAGATCTTAATGTAAGTGAACAACAATGTTGATGCAACTGTTTCAATTA 2992
 Qy 6892 AATTAATTTGCAATTAAGTGAAGAAACAGATATGATCATAGTACTTTTCACTTTCTACT 6951

Db 2993 AAAGATAGTCAAAATTAATTCATTAATCTCTTAACATTAAGTGAAGAACTGCTAATACAGCT 3052
 Qy 6952 GATGTAATATGTAAGAGCTCTTAATTAATAATTTGAAATTTCTGACAGCCGCTGAGAGACC 7011
 Db 3053 GTATTTGCTAACACTGTTTCAAGCTGTGATACATTAATCTTTTAACTGCTGTGATCAATTA 3112
 Qy 7012 GCAGGCTTGGACAGATTAACCGAGTGTGTTCTGTTAACAATTAATAGTTCTGTGATA 7071
 Db 3113 GTTACTTATGCAAGATGCTTAAATAATGCTGACAGTGTGCTGAAATAATTAATCACTGACGCTA 3172
 Qy 7072 GCTGAGTTCACAAATTAATCTGATTTGATCTCCGTAAGAGAAATGTAATGTAACGCA 7131
 Db 3173 ACATTAAGAAATTAATCTGACAGCAATTAATCTGATTAATTAACAGAGTGTATTAACA 3232
 Qy 7132 AAAGAGCAAAAATTAATTAAGCAAAACAGCAAAATGCAAGATTCGAGAGACAGCAATC 7191
 Db 3233 TCACAGCTACAGAGCTGAATATATCTTC-----TAAATCAATTTGCTGCAAGATTAATCA 3286
 Qy 7192 GGAGCCAAATGCTTGTGTAATTAATTTTGGAAACAGCTGTAGAAATGAAAAATTTCTGAA 7251
 Db 3287 TTGCAACAGGTGAAGATTCACCTTAATTAATGATTAATGCTGTGCTCAAGTAAATTAAC 3346
 Qy 7252 GGAAGAGCAAGAGATTTTAAATCTTTAGACGAATTAACAAGAAACAAAGATTAATAA 7311
 Db 3347 TTACAGGTAAAGAAAGGTGCAAGGTGTAGCTGATGATCAATGCTACATTTGACAGT 3406
 Qy 7312 GTAATGATGCTTACGAAATAATCTTAATCAATCAGCAGATTTTCTACAGAGATTAATCTCT 7371
 Db 3407 ACTGCAACTGTTTCTGAGAGCAAAAGTATTAATTAATCAGTACAAACAGTGTGTTCT 3466
 Qy 7372 GTAAAGCGATGAGAGATTAATCTCAGAGAGAGAAATTAAGCAATTTGTAAGATTTCT 7431
 Db 3467 GAAATGGAATTAATCTCTCTGTTAATCAAGTATTAATGCAAGTAAAGTTAACGGTAA 3526
 Qy 7432 GATATTAATGCAAAAATGATGATTAATTAACAACAGAGCAAGAAATTAATTAATCACTTACT 7491
 Db 3527 GATCAAGCTGTGCAAGAAACGCTGTACAAACAAATTCACAGATTAATCAAGCCCTTCT 3586
 Qy 7492 GGTGATTTGGGAACTGCAGAGCTTCTGCTCCGATCAGAAACAGTGGCAGTTACAAATAT 7551
 Db 3587 GTGGGTGAAAAAGTATGATTAATGATGATGATTAATTAATGCTGTAGCAATTTGGAACGCT 3646
 Qy 7552 AAAAGAAATTCGAGATTAATCTGTTGAATAATCTTTTGGAAAGCAGTGAATAAT 7611
 Db 3647 CCAACAGCAAAATTAATCTGATGATTAATGATGATTAATTAATTAATGATTAATGATTA 3706
 Qy 7612 GTTATGATGATTAATTAACAGAAATGTTGCTTTAACAAGATTAATCAAGTCCCTGTAGAGCA 7671
 Db 3707 CAAGCTGCAAAATCTTGTCTGCTACAAATGATTAATTAACAGTCAAGATTAATTAACAGCT 3766
 Qy 7672 TTGGAATAGAGAGCTGCTATGCAAGATTAATTTCTAATGGAAGATCAATATCAAGTAT 7731
 Db 3767 TCTGCAACAGGTGCTACATTAATC---ATTAATCTTACTGTGAACAACAGTATGATTAACA 3823
 Qy 7732 AAAAATTTAACTAATTAAGAAAAAATTAATGATGATTAATTTGTAAGA 7778
 Db 3824 ATTAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3870

RESULT 30
 US-08-682-517-8
 ; Sequence 8, Application US/08682517
 ; Patent No. 5874267
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: Expression of surface layer proteins
 ; NUMBER OF SEQUENCES: 25
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/682,517
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 4197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Bacillus sphaericus
INDIVIDUAL ISOLATE: P-1
FEATURE:
NAME/KEY: CDS
LOCATION: 95..3850
FEATURE: NAME/KEY: mat_peptide
LOCATION: 185..3850
FEATURE: NAME/KEY: sig_peptide
LOCATION: 95..184
US-08-682-517-8
Query Match 0.6%; Score 59.4; DB 2; Length 4197;
Best Local Similarity 42.0%; Pred. No. 0.03;
Matches 826; Conservative 0; Mismatches 1111; Indels 30; Gaps 7;
QY 5830 ATAGAGATTAAATCATGCAAAATTCCTGGAAGGGAATGTGGAGTTATTAACAG 5889
DB 1916 AAAGCAAAACAATCTCTGTAAACAAATGATTAAGGTGAGCTACATTTTTCGT 1975
QY 5890 TCTGATGCGGTAAATGCTAATTATGACAGAACAGTGTCTGAGTGGCCGTCAGCAATA 5949
DB 1976 TCTGATGAGTAAAGATTATGCAACCAATTTGCTGATTTGATTAATTAATCTTGAT 2035
QY 5950 GGAGCCTCAACAGTGTGATTAATAATACAGATTTCAAAAGCATATGTAAAGATTCT 6009
DB 2036 GCAAAACAAGCGACTTGATGAGTGAACCAAAAGCAGTTCACCAATCTTACTTC 2095
QY 6010 ACAGGATTCCTAAAGAAAGAACAGATTAATTAATTAATTAATTAATTAATTAATTA 6069
DB 2096 CAAGCAACATATCTTGATGCTCAGCTATCAAGCATACAAAATTCAGATCTTAATTA 2155
QY 6070 GTGTAGATTAAAGTATTCAAAATCTTAATTAATTAATTAATTAATTAATTAATTA 6129
DB 2156 GCTGATACCTAATTTGATGCTTCTGAACTGCAATTTTGACAGCAATTAATTAACCA 2215
QY 6130 ATTAAGTAAATAAAGATTTGTTACCAATAGTTCAGTACTCATCTTTAAATCTTTA 6189
DB 2216 ACCGCAAAAAGTAACGTGTAATCTTAAGAAAGCAACTTATTAATTAATTAATTA 2275
QY 6190 TTGGCAAAATCCGCTGCTTCAAGACAGCCGAGTGGCAGAACTGTTAATTAATCA 6249
DB 2276 GGTGCTAATGATTAATTAAGTAACTGTAACCAAGTTATCTCAAAATGTGCTACAG 2335
QY 6250 GTTTATGAGAAACAGAGCTCTTGTAGAAAATTTCTATATTAATGCAAAACATTAAT 6309
DB 2336 ACTTATGAAAGCTACTTATCTTCTACAGAACTGTTATTAACCTGCTAAGAAATTA 2395
QY 6310 GTTAAATCAAGAGATTACAGAAATTCATTCGAGTATGATGTTCTGTTGGTGTGGA 6369
DB 2396 GTTACTTCAGTGTGATTAACCACTGCTTAAAGTAAATGCTACAGGTAATGCTGTT 2455
QY 6370 AATGTGAGATGAGAGCTTCTTGATTAACCAATATTAATTAATTAATTAATTAATTA 6429
DB 2456 AATACAGACGTAAAGACTATGCAATTAATGCTAAAGAACTACACTACATTCACAGCT 2515
QY 6430 GTTGAAGAAATCAATGCTGTGATGAAGTTTCGAGAAAGCTGAAATTAACAGAGAT 6489
DB 2516 ACAAATGAAGTTCCAACTTTTACACGTGTGATGCTACTCA---TTCAATACAGCTGAT 2572
QY 6490 TCTAAGCAAGAAATTTCTCTTTTGGAGTGGAGTGCAGACAGCCGGGATGAGAGCCGA 6549

DB 2573 TCTGTTCAACACCACTATTTGTTTGTGTAATAAACCAGGAAATATGCTGCT 2632
QY 6550 GTGCGAGAAACCGTTTCGTAAATCAATTTGACAGAAAGACGAGTATGATGGAAGA 6609
DB 2633 GTATCAGGCAAAACATTAATTAATCTTGGAGCTAATGCTATATGATTAATTTGGAGCG 2692
QY 6610 GCA---AAGATTTTGGTAAATAAAGCTGAGATTACAGCAAAAGCTTATAGTCTGTTGC 6665
DB 2693 GCATGGAAAGCATTAATTAATCAATATGCACTAAGGCGCAAAAGTAAATCAATCATAT 2752
QY 6666 AA---TTGAAAATGCGCAGTGGAGTGGCTGCAAAAGAGCTGAAATTTGAGCAGAGTG 6723
DB 2753 AATGTAATGCTGATACAGTTACATTTAAAGTAATTAATGCTGTAATTTCTCACTGA 2812
QY 6724 GCAGTTACCAAGATGATCAACACAGACAGCAAGAGTGAATAATTC---TAAATTA 6778
DB 2813 GCTATCAAAACAGTTGCTCCAAACAACAGCACTCCAACTAGCGGCATTAACATTA 2872
QY 6779 TGACTCGAAACAGT-----TAGATGTAATACCAAAATGATTAATTAATCAAGTACT 6831
DB 2873 ACACCAAGACGTGTGTTAGTTGATTTAACTCACTCACTCACTTAACATTTGAAATTTCA 2932
QY 6832 GGAATCGGTTGAGCCGAGCTGGAATTTCTGACCGGAGTATCTGAGTGTTCCTGTC 6891
DB 2933 TTAGCTGATGAGATCTTAATATGTAAGTCAACCACTGTTGATCTGCACTGTTTCATTA 2992
QY 6892 AATTAATTTGCAAAATGATGAAGAAACAGATATGATATGATCTTAACCTCTTACT 6951
DB 2993 AAAGATATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3052
QY 6952 GATGTAATGTAAGAGCTCTTAATTAATTTTCGAATTCCTTGACACCGCGGAGAGACC 7011
DB 3053 GTATTGCTACACCTGTTCAAGCTGTGATCAATTAATCTTTTAACTGCTGATCAATTA 3112
QY 7012 GCAAGTCTTGACAGAGTTACCGAGTGTGTTCTGTTAACAATTAATTAATTAATTAATTA 7071
DB 3113 GTTACTTAATGAGATGCTAAATAATGCTGAGTGTGCTGAAATTAATTAATTAATTA 3172
QY 7072 GCTCAGTTTCACAATTAATCTGATTTGAATTCCTGTAACAGAAAGTAATTAATTAATTA 7131
DB 3173 ACATTAAGAAACATTAAGTGAAGCAATTAATCTTGTATACATTAACAGAGTATTAACA 3232
QY 7132 AAAGGAAATAAATTAATTAAGCAACAGACGAAATGCAAGATTTGGAGAGCAGCAATC 7191
DB 3233 TCAGCAGCTACAGCAGCTGAATTAATCTTC-----TAAATCAATGCTGCAAGATTATCA 3286
QY 7192 GGAGCAATGCTTGTGAATTAATTTTGAACAGCTGTAGAAATGAATAAATTTCTGA 7251
DB 3287 TTGCAACAGGTGAAGATTTCACTTTTAATTAATTAATTAATTAATTAATTAATTAATTA 3346
QY 7252 GGAAGAGAAACAGAGTTTAAATACTTAAAGCAGAGTTAAACAAAGAACAGATTAATAA 7311
DB 3347 TTAGCAGGTAAATAAAGTGAACCAAGGTGATGATGATTAATTAATTAATTAATTAATTA 3406
QY 7312 GTAATGATGCTACGAAATAAATCTTAATTAATGAGAGTATTTCTACAGAAATCTTCT 7371
DB 3407 ACTGCACTGTTTCTGAGACCAAGTATGTTAATCAAGCTAACAAGGTGTTGTTCT 3466
QY 7372 GTPAAAAGGAGATGAGAGATPACTCAGGAGAAAGAAATTAAGCATTGTGAAGCTTCT 7431
DB 3467 GAAAGTTGAATTAATTAATCTTCTTCTTGAATTAATTAATTAATTAATTAATTAATTA 3526
QY 7432 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3586
DB 3527 GATCAAGTGTGTCAGAGAACAGCTCTCAAAAGCATTCACGATTAATTAATTAATTAATTA 3646
QY 7492 GGTGTTTGGAACTGCAAGTCTTCTGCTTCCGATACAGAAACAGTGTGCAATTAATTA 7551
DB 3587 GTGGGTGAAAGAAATGATTAATGATGTTGATTAATTAATTAATTAATTAATTAATTAATTA 3646
QY 7552 AAAAGAAATTCGAGATTAATGTTGAAATTTCTTTTGTGAAGAGCTGAAATAATTAAT 7611

QY 7989 ATCTCAGAGCTTCTGTAGAGCAGTACAGGGCCAGGAATTTTCCGAACAAAAGA 8048
DB 2387 AGTTAGAGAGCAGAGCAGAGGAGTTAGAGCAGAGCAGAGGATTAGAGGCGAGAGC 2446
QY 8049 TGC CGAAGCTCTTATTGAAAAGTTAGTACAAAATCCCGAAGAACTATTTTTCATGCGA 8108
DB 2447 AGGAGTTAGAGAGCAGAGCAGAGGAGTTAGAGAGCAGAGCAGAGGATTAGAGAGCAGG 2506
QY 8109 TAATGTGATATGAGAACACACATAAATGAAATGAAACAGAGCTTTCTAAAGCAGTAAAC 8168
DB 2507 AGCAGAGTTAGAGAGCAGAGGAGTTAGAGAGCAGAGCAGAGGAGTTAGAGAGCAGAGC 2566
QY 8169 AGGTTCTGTATTGGAGAGAGTTGAGTCAACAGGAGAGAGTACTGCTCAGTAAAC 8228
DB 2567 AGGAGCAGAGAGCAGAGGAGTTAGAGAGGTTGAGAGCAGAGCAGAGCAGAGGAGC 2626
QY 8229 TATGTTAGAGTTGAGAGAGAAATTTGTCGAAACAAATGCAATTCATTTCTTA 8288
DB 2627 AGGAGAGCAGAGGAGTTAGAGAGGTTGAGAGCAGAGAGAGCAGAGGTTAGAGAGGTTG 2686
QY 8289 AGTAAAGGTTTGTGATGAAGATTAAGTAACTGCTAAATCTTCTGTAGTATCAGAAATGG 8348
DB 2687 AAGAGCAGAGAGAGCAGAGGATTAGAGAGGTTGAGAGCAGAGCAGAGGATTAGAGC 2746
QY 8349 AGGAGAAATTCGCGAGCAGAGGAGTGAATCTTCTACAGCAGCAAAAGTAACTAGAA 8403
DB 2747 AGGTGAAGAGCAGAGCAGAGGAGGAGTGAACAGAGAGCAGAGCAGAGCAGGTTGA 2801

RESULT 35

US-09-894-273-1
Sequence 1, Application US/09894273
Patent No. 6756203
GENERAL INFORMATION:
APPLICANT: Kieff, Elliott D.
APPLICANT: Ballesteras, Mary E.
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/894,273
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/109,422
PRIOR FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3489
TYPE: DNA
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-1

Query Match 0.6%; Score 59; DB 3; Length 3489;
Best Local Similarity 45.3%; Pred. No. 0.034;

Matches 215; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 7929 TAAAGCAATTTGAGAGAAATCAATGTCAAAAGTGAAGAAAAGAGAGTACTGCTGA 7988
DB 2327 TAGAGAGTACAGAGCAGAGGATTAGAGAGCAGAGCAGAGGATTAGAGAGCAGAGCAGG 2386
QY 7989 ATCTCAGAGCTTCTGTAGAGCAGTACAGGGCCAGGAATTTTCCGAACAAAAGA 8048
DB 2387 AGTTAGAGAGCAGAGCAGAGGAGTTAGAGAGCAGAGCAGAGGATTAGAGGCGAGAGC 2446
QY 8049 TGC CGAAGCTCTTATTGAAAAGTTAGTACAAAATCCCGAAGAACTATTTTTCATGCGA 8108
DB 2447 AGGAGTTAGAGAGCAGAGCAGAGGAGTTAGAGAGCAGAGCAGAGGATTAGAGAGCAGG 2506
QY 8109 TAATGTGATATGAGAACACACATAAATGAAATGAAACAGAGCTTTCTAAAGCAGTAAAC 8168
DB 2507 AGCAGAGTTAGAGAGCAGAGGAGTTAGAGAGCAGAGCAGAGGAGTTAGAGAGCAGAGC 2566
QY 8169 AGGTTCTGTATTGGAGAGAGTTGAGTCAACAGGAGAGAGTACTGCTCAGTAAAC 8228

DB 2567 AGGAGCAGAGAGCAGAGAAATTTAGAGAGGTTGAGAGCAGAGCAGAGCAGAGAGC 2626
QY 8229 TATGTTAGAGTTGAGAGAGAAATTTGTCGAAACAAATGATTTGAAATGCAATTTCTTA 8288
DB 2627 AGGAGAGCAGAGGATTAGAGAGGTTGAGAGCAGAGAGCAGAGGATTAGAGAGGTTG 2686
QY 8289 AGTAAAGGTTTGTGATGAAGATTAAGTAACTGCTAAATCTTCTGTAGTATCAGAAATGG 8348
DB 2687 AAGAGCAGAGAGAGCAGAGGATTAGAGAGGTTGAGAGCAGAGCAGAGCAGAGGATTAGAG 2746
QY 8349 AGGAGAAATTCGCGAGCAGAGGAGTGAATCTTCTACAGCAGCAAAAGTAACTAGAA 8403
DB 2747 AGGTGAAGAGCAGAGCAGAGGAGGAGTGAACAGAGCAGAGCAGAGCAGGTTGA 2801

RESULT 36

US-08-770-379-20/C
Sequence 20, Application US/08770379
Patent No. 5849564
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESS: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/770,379

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 52342

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 3207 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-770-379-20

Query Match 0.6%; Score 59; DB 2; Length 3207;
Best Local Similarity 45.3%; Pred. No. 0.07;

Matches 215; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 7929 TAAAGCAATTTGAGAGAAATCAATGTCAAAAGTGAAGAAAAGAGAGTACTGCTGA 7988
DB 19670 TAGAGAGTACAGAGCAGAGGATTAGAGAGCAGAGCAGAGGATTAGAGAGCAGAGCAGG 19611
QY 7989 ATCTCAGAGCTTCTGTAGAGCAGTACAGGGCCAGGAATTTTCCGAAGCAAAAAGA 8048
DB 19610 AGTTAGAGAGCAGAGCAGAGGATTAGAGAGCAGAGCAGAGGATTAGAGAGCAGAGC 19551

QY 8049 TGCCGGAAGCTCTTATTGAAAGTTAGTACAAATCCGGAAGACTATTTTCATGCAGA 8108
DB 19550 AGGACTTGTAGAGACAGACAGAGTTTAGAGACAGAGAGAGTTAGAGAGCAGG 19491
QY 8109 TAATGTGAATATGGAACAACATTAATGAATGAACAGAGTTTCTAAAGCAGTAAC 8168
DB 19490 AGCAGAGTTAGAGAGCAGAGGTTGAAAGCAGACAGAGAGGTGAGAGCAGAGC 19431
QY 8169 AGCTTCTGTATTGGAGAGAGTTGGAGTCAACAGGAGAGAGTACTGCTGCAGGTAAC 8228
DB 19430 AGGACAGAGAGAGAGAGGAAATTTAGAGAGTGAAGAGAGAGAGAGAGAGC 19371
QY 8229 TATGTAGAGTTGAGAGAGAAATTTGTCAAGCAATGATGAAATGCAATTTCTTA 8288
DB 19370 AGGAGAGACAGAGTTAGAGAGGTTGAAAGAGAGAGAGAGAGTTAGAGAGTGG 19311
QY 8289 AGTAGAAGCTTTGATGAATGAATGAATGCTTAATCTTCTGTAGTATCAGAAATGG 8348
DB 19310 AAGACAGAGAGAGAGAGAGTTAGAGAGGTTGAAAGAGAGAGAGAGAGGTTAGAG 19251
QY 8349 AGGAGAAATTCGCGAGCAGAGAGTGAATATCTTACAGACCAAGTAACTGAA 8403
DB 19250 AGGTGAAGAGAGAGAGAGAGGAGGAGTGAACAGCAGAGAGAGAGAGCGGTGA 19196

RESULT 37
US-08-757-669A-20/c
Sequence 20, Application US/08757669A
Patent No. 6183751
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,669A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45185-F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 32207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-669A-20

Query Match 0.6%; Score 59; DB 3; Length 32207;
Best Local Similarity 45.3%; Pred. No. 0.07;
Matches 215; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 7929 TAAAGAAATTGAGAGAGAAATCAATGTCAAAGTGAAGAAAAAGAGAGTGAAGTCTGA 7988
DB 19670 TAGAGATCAGAGACAGAGAGTTAGAGAGCAGAGAGAGAGTTAGAGAGCAGAGCAGG 19611
QY 7989 ATCTCAAGAGCTTCTGTAGAGAGAGTGAAGAGGCGAGAGAAATTAATTCGAGCAAAAGA 8048
DB 19610 AGTTAGAGAGCAGAGAGAGAGTTAGAGAGCAGAGAGCAGAGAGTTAGAGAGCAGAGC 19551
QY 8049 TGCCGGAAGCTCTTATTGAAAGTTAGTACAAATCCGGAAGAGATATTTTCATGCAGA 8108
DB 19550 AGGAGTTAGAGAGCAGAGAGAGTTAGAGAGCAGAGAGAGAGTTAGAGAGAGCAGG 19491
QY 8109 TAATGTGAATATGGAACAACATTAATGAATGAACAGAGTTTCTAAAGCAGTAAC 8168
DB 19490 AGCAGAGTTAGAGAGCAGAGGTTGAAAGCAGACAGAGAGTGAAGAGCAGAGC 19431
QY 8169 AGCTTCTGTATTGGAGAGAGTTGAGTCAACAGGAGAGAGTACTGCTGCAGGTAAC 8228
DB 19430 AGGACAGAGAGAGAGAGGAAATTTAGAGAGTGAAGAGAGAGAGAGAGAGCAGG 19371
QY 8229 TATGTAGAGTTGAGAGAGAAATTTGTCAAGCAATGATGAAATGCAATTTCTTA 8288
DB 19370 AGGAGAGCAGAGTTAGAGAGGTTGAAAGAGAGAGAGAGAGTTAGAGAGTGG 19311
QY 8289 AGTAGAAGCTTTGATGAATGAATGAATGCTTAATCTTCTGTAGTATCAGAAATGG 8348
DB 19310 AAGACAGAGAGAGAGAGTTAGAGAGGTTGAAAGAGAGAGAGAGAGGTTAGAG 19251
QY 8349 AGGAGAAATTCGCGAGCAGAGAGTGAATATCTTACAGACCAAGTAACTGAA 8403
DB 19250 AGGTGAAGAGAGAGAGAGGAGGAGTGAACAGCAGAGAGAGAGCGGTGA 19196

RESULT 38
US-09-230-371A-20/c
Sequence 20, Application US/09230371A
Patent No. 6348586
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 45185-G-PCT-US
CURRENT APPLICATION NUMBER: US/09/230,371A
PRIOR FILING DATE: 1999-11-17
CURRENT APPLICATION NUMBER: PCT/US97/13346
PRIOR FILING DATE: 1997-07-22
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 20
LENGTH: 32207
TYPE: DNA
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20

Query Match 0.6%; Score 59; DB 3; Length 32207;
Best Local Similarity 45.3%; Pred. No. 0.07;
Matches 215; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 7929 TAAAGAAATTGAGAGAGAAATCAATGTCAAAGTGAAGAAAAAGAGAGTGAAGTCTGA 7988
DB 19670 TAGAGATCAGAGACAGAGAGTTAGAGAGCAGAGAGAGAGTTAGAGAGCAGAGCAGG 19611
QY 7989 ATCTCAAGAGCTTCTGTAGAGAGAGTGAAGAGGCGAGAGAAATTAATTCGAGCAAAAGA 8048
DB 19610 AGTTAGAGAGCAGAGAGAGAGTTAGAGAGCAGAGAGCAGAGAGTTAGAGAGCAGAGC 19551
QY 8049 TGCCGGAAGCTCTTATTGAAAGTTAGTACAAATCCGGAAGAGATATTTTCATGCAGA 8108

Db 19550 AGGAGTTAGAGGACGAGGACGAGGATTAGAGACGAGGAGGATTAGAGACGAG 19491
Qy 8109 TAAATGTAATATGAGGACCAATATAATGAAATGAAACGAGTTCTTAAAGCAGTAC 8168
Db 19490 AGCAGAGTTAGAGAGCGAGGAGTGGAAAGCAAGAGCAGAGAGGTGGAGAGCAGAGC 19431
Qy 8169 AGGTTCTGATTGGAGAGGAGTTGAGTCACCAAGCAGAGAGTACTGTCAGGTTAAAC 8228
Db 19430 AGGACGAGGAAAGACGAGAAATTAGAGAGGTGGAGAGAGCAAGAGGAGGAGAGGAGC 19371
Qy 8229 TATGTTAGAGGTTGAGAGGAAATTTGTTCAACCAATTCATGTAATGCAATTTCTTA 8288
Db 19370 AGGAGAGGAGGAGGATTAGAGAGGTGGAAAGCAGAGGAAAGAGCAGGATTAGAGAGGTGG 19311
Qy 8289 AGTAAAGGTTTGAATGAAATGAAATGAACTGCTAAATCTCTGATGATCAGAAATGG 8348
Db 19310 AAGACGAGGAAAGCAGGAGGTTAGAGAGGTGGAAAGAGCAGAGAGCAGGAGTTAGAGG 19251
Qy 8349 AGGAGGAAATGGCCGAGCAGAGAGTGAATCTTCTACAGACCAAAAGTAACTAGAA 8403
Db 19250 AGGTGAAGAGCAGGAGCAGAGGAGGTGGAAACAGACAGGAGCAGGAGAGCGGTGA 19196

RESULT 39
US-08-682-517-13
; Sequence 13, Application US/08682517
; Patent No. 5874267
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Expression of surface layer proteins
; NUMBER OF SEQUENCES: 25
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/682,517
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3666 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-682-517-13

Query Match 0.6%; Score 58.2; DB 2; Length 3666;
Best Local Similarity 41.7%; Pred. No. 0.05;
Matches 801; Conservative 0; Mismatches 1093; Indels 27; Gaps 6;

Qy 5830 ATAGCAGAAATTAATCATGCAAAAATCACTGCGAAGGAAATGTCGAGTTATTAACAGAG 5889
Db 1732 AAAAGAAAACAATCTCTGTAATAAACAATGATTAAGGTAAGCAATTTGTTATCGGT 1791
Qy 5890 TCTGATGCGTAAATGCTAATTAATGACAGAACAGTGTCTGAGTGGCCGTCAGCAATA 5949
Db 1792 TCTGATACAGTAAGATTAATGCAACCAATTTGCTGATGATTAATTAATCTTCTGAT 1851
Qy 5950 GGAGGCTCAACAGGTGTGAATGAATTAACAGGATCTCAAAAAGCAATATGTAAGAAATCT 6009
Db 1852 GCAAAACAAGGAGCACTTATGTAAGTGAACCAAAAGCAGTGGCAACATCTCTTACTTC 1911
Qy 6010 ACAGATGATTGCTAAAGAAAGAAACAGATGATTAATTAATCAAGGCGAGTAGATTA 6069
Db 1912 CAAGCAACCATATCTGATGCTCAGCTATCAAGACATACAAAAATCAGATCTTAATAA 1971
Qy 6070 GTGTGATGATTAAGTATTCAAAAATCTTAATATTAACGAAGACTTATCACAAAAAGAAA 6129
Db 1972 GCTGTAACTAAAGTTGATGAGTTCTGAAACATGCAAGTATTTGACAGAAATTAAGTAAACCA 2031

Qy 6130 ATAAATATAAAAAAGATTTGTTACCAATAGTTCACTACTCATTTAAAACTTTA 6189
Db 2032 AGCGGCAAAAAAGTAAGTACGTACTTTATTAAGAAAGCACTTATTAACAATCTCAACT 2091
Qy 6190 TTGGCAATGCCGCTGGTTCAAGACAGCCGAGTGGAGGAACTGTTAATATCAACAG 6249
Db 2092 GGTCTATATGATTAATTAAGTATGATTAACCAATTTATCTCAACMAATCGTAGTACACAGTA 2151
Qy 6250 GTTTATGAGAAACAGAAAGCTCTGTGAAAAATCTAATTAATTAATGCAAAACATTAATCT 6309
Db 2152 ACTTATGAGAGTACTTTATCTTACAGAACTGTTATTAACCTGCTCAAAATTTAGAA 2211
Qy 6310 GTAAATCAGAGATTAACCAATTCATCCGAGTAGTAGTCTGTTGGTGGTGA 6369
Db 2212 GTTACTTACAGTAGTATGTAAGAAACAGCTGTTAAAGTAATGCTACAGATATGCTGT 2271
Qy 6370 AATGAGAGTAGAGAGCTTCTTGATATCAATTTATTAAGAAATGCAACAGCAAGA 6429
Db 2272 AATACAGACGTTAAAGACTATGCAATTAAGTAAAGAGCTACAGCTACATTCACAGCT 2331
Qy 6430 GTTGGAAAACTACAAATGCTGATGAAGGTTCCGAGAAAGAGCTGAAATTAACAGCAGAT 6489
Db 2332 ACAAATGAAGTTCCAAACTCTTACACTGATGATGCTACTCAA--TTCAAATACAGCTGAT 2388
Qy 6490 TCTAAGCAAGAAATTTCTCTTTTGAAGTCCGAGTCCGAGCAGCCGAGGTAGAGCCGGA 6549
Db 2389 TCTGTTCAACACCAACTATTTGTTGTTGCTGTAAGAAACCAAGTAATATGCTGT 2448
Qy 6550 GTGCGAGAACCGTTTCCGTAAATCAATTTGACAGAAAGACGAGAGTAGATGAGAGAA 6609
Db 2449 GTATCAGCAAAACATATAAATTAATCTCCGAGCTAATGTAATGAAGTATTTGGTAAGCG 2508
Qy 6610 GCA----AAGATTTTGGTAAGAAAGCTGATTAACGCAAAACGTTAATTTCTGTTGC 6665
Db 2509 GCATGGGAAGATTAATTAATCAATATGCAATGCAACTGAGGCCAAAAGTAACAATTCATAT 2568
Qy 6666 AA--TTGGAATGCGCGAGTCCGAGTGGCTCAAAAGAGCTGGAATTTGAGCAGCAGTG 6723
Db 2569 AATGTAGATGATGATACGATTAATTAAGTAAATTAATGATGCTGTTAATCTTCAACTGAA 2628
Qy 6724 GCAATTACCAAAAGATGAATCAACACGAGACCAAGTGAATAAT--TCTAAATTA 6778
Db 2629 GCTATCAAAACGATGCTCCAAACACACAGCTCCCACTAGCGGCATTAACATTA 2688
Qy 6779 TGACTCGAAGCAAGT-----TAGATGTAATAGCAGAAATGATTAATTCAGGTACT 6831
Db 2689 ACACGAGCAGCTGGTGGTTAGTTGATTAACAATGCAACTACATTAAGAAATTTCA 2748
Qy 6832 GAAATCGGTTACGCCGAGCTGGAATTTCTGACGCCGAGTATCTGAGAGTGTCTGTC 6891
Db 2749 TTAGCTGATGCAAGATCTTAATGTAAGTCAACAACGTGTTGATCTGCAACTGTTCAATTA 2808
Qy 6892 AATAATATTCGCAATTAAGTGAAGAAACAGATATCGATCATTAATCTTCACTCTTACT 6951
Db 2809 AAAAGTATGCAAAATTAATCTTATCTCTTACATTAAGTGAACGTGCTAATACAGGT 2868
Qy 6952 GATGTAATGTAAGACCTTATTAATAATTCGAATTCCTGACAGCCGCGAGAGAGCC 7011
Db 2869 GTATTTGTACAACTGTTCAAGCTGTGATCATTAATCTTTTAACCTGCTGTATCAATTAACA 2928
Qy 7012 GCAGGCTTTCGAGAGATTAACGAGAGTGGTTCTGTTAACATTAATTAATAGTTCTGTGATA 7071
Db 2929 GTTACTTATGCAAGATGCTTAAGAAATGCTGAGGTGTGTAAGAAATTAATCTGTAAGCTTA 2988
Qy 7072 GCTGAGTTACAAATTAACCTGATTTGATTTCCGTACAGAGAAAGTAATGTAACGCA 7131
Db 2989 ACATTAAGAAAGAACTACTGAGAGCAATTAATCTGATACATTAACAGAGTATTAACA 3048
Qy 7132 AAAAGGAAAAAATATTAACCAACAGACCAATGAGAGATTCGAGAGCAGCAATC 7191
Db 3049 TCAGCAGCTACAGCAGCTGAATATATCTTC-----TAATCAATTTGCTGAGATTAACA 3102
Qy 7192 GAGCCAAATGTCTTGATTAATTAATTTTGGAAACAGCTGTAGAGAAATGAAAAATTTCTGA 7251

Db 3103 TTTCACACAGGTGAGGATTCACCTTTAAATATGATTAATGCTGCTCAAGTAATTAC 3162
Qy 7252 GGAAGAAGACAGAAAGTTTAAAACTTTAGACGAGTAACAAAGAAAGATTAAGAA 7311
Db 3163 TTACGAGGTAAAGAGAGTGCACAGGTGTACGTATGCTATCAATGCTACATTTGCGAGT 3222
Qy 7312 GTAATGATGCTACGAAAAAATCTTACAAATCAGAGGTAATTTTACAGAAATATCTTCT 7371
Db 3223 ACTGCACTGTTTCTGAGACAAAGTACTTAAATCAGCTACAAACAGGTGTTGTTCT 3282
Qy 7372 GTAAAAAGCGATAGAGAGTACTCAGGAGAGAAATTAAGCCATTGTGAAGACTTCT 7431
Db 3283 GAAGTTGAAGTTACATCTCTCTCTGTTAATCAAGATTAATGACAGTGAACGGTAA 3342
Qy 7432 GATATTAATGAAAAATGATGATATTAACAAGAGACAAAGATTAATCAGCTTCTACT 7491
Db 3343 GATCAAGTCTTGACAGAAACAGCTCTACAAAAGCATTCAGATTAATCAAGCCCTTCT 3402
Qy 7492 GGTGTTTGGAACTGACAGTCTTGTCCGATCAGAAACAGTGCAGTTACAAATAT 7551
Db 3403 GTGGGTGAAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3462
Qy 7552 AAAAGAAATCCGAGTACTGTTGAAAATTTTGTGAAAAGCAGCTGAAAAAGTAAAT 7611
Db 3463 CCAACAGCAAAATACATTCGTAGTGAATCTGCTGCTAAATACATTCGTAGTAC 3522
Qy 7612 GTTAGATGGAATATTAACAGAAATGTTGCTTAAACAGCATACAGGCTCTGAGAGCA 7671
Db 3523 CAGCTGCAAACTTGTGCTACAAATGATCTTTAAACATGCAAGATTAAGTTACAGCT 3582
Qy 7672 TTGGAATAGAGCTGCTATGACAGAAATTAATCTAAATGGAAGATCAAAATTCAGTAT 7731
Db 3583 TCTGCAACAGGTGCTACTATTAATTAATCTTACTGTAACACCAAGTGTGCTACAAAT 3642
Qy 7732 A 7732
Db 3643 A 3643

RESULT 40
US-08-682-517-14
; Sequence 14, Application US/08682517
; Patent No. 5874267
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Expression of surface layer proteins
; NUMBER OF SEQUENCES: 25
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,517
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3666 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3666
US-08-682-517-14

Query Match 0.5%; Score 58.2; DB 2; Length 3666;
Best Local Similarity 41.7%; Pred. No. 0.05;
Matches 801; Conservative 0; Mismatches 1093; Indels 27; Gaps 6;

Qy 5830 ATAGCAGAAATTAATCAATGCAAAAAATCACTGCCAAGGGAATGTCGAGATTAACAGAG 5889
Db 1732 AAAGCAAAACAAATCTCTGTAAGAAACAAATGATTAAGGTGAAGCTACATTTGTTATCGGT 1791
Qy 5890 TCTGATCCGTAATTTGCTAAATTAATGACAGAAACAGTGTCTGGAGTGGCCGTCAGCAATA 5949
Db 1792 TCTGATACAGTAACAGATTAATGCAACCAATGCTGATGATTAATTAATTAATCTTCAT 1851
Qy 5950 GGAAGCTCAACCAAGTGAATGAATTAACAGATCTACAAAAGCATATGTAAGAAATCTCT 6009
Db 1852 GCAAAAACAGGCGACTTGAATGAAGGTGAACCAAAAGCAGTTGACCAATCTCTTACTTC 1911
Qy 6010 ACAGTATGCTAAAGAAAGAAACAGATGATTAATTACTACTCAGGCGAATGATGAATA 6069
Db 1912 CAAACACCAATCTTGAATGCTGCTACAGTATCAAGCATACAAAATCAGATCTTAATTA 1971
Qy 6070 GTGATGATTAAGATTAATCAAAAATCTTAATTAATTAACGAACTTAATCAAAAAAGAAA 6129
Db 1972 GCTGTAATTAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2031
Qy 6130 ATAAATTAATTAAGAAATGATTTGTTACAAATGCTTCACTACTACTTAAATCTTTA 6189
Db 2032 AGCGCAAAAAGTAAGTGAATGATCTTATTAAGAAAGCACTTAATCAATCTACAAATCT 2091
Qy 6190 TTGCAAAATGCCGCTGCTGATCAGAACAGCGGAGTGGCAGGAAGCTTAATATCAACAG 6249
Db 2092 GGTGCTAATGATTAATTAAGTGAATACCAATTAATCTCAAAATGTAAGTACTACAGTA 2151
Qy 6250 GTTATGAGAAACGAAAGCTCTGTGAGAAATCTATTAATTAATTAATTAATTAATTAAT 6309
Db 2152 ACTTATGAAGTACTTATTAATCTTACAGAACTGATTAATTAATTAATTAATTAATTAAT 2211
Qy 6310 GTAATATCAGAGATTAACAGCAATTAATCAAGTGAATGATGATGATGATGATGATGAT 6369
Db 2212 GTTACTTCAAGTGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2271
Qy 6370 AATGTAAGATGAAGAGCTTCTGATACCAATTAATTAATTAATTAATTAATTAATTAAT 6429
Db 2272 AATACAGACGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2331
Qy 6430 GTTGAAGAAACATCAATGCTGATGAAGTTCGAGAAAGCTGAATTAATTAATTAATTAAT 6489
Db 2332 ACAATTAAGATTCAAATCTTACACTGTGTGATGATGATGATGATGATGATGATGATGAT 2388
Qy 6490 TCTAAGCAAGAAATTCCTCTTTTGAAGTCCGAGTCCAGACAGCCGGGTAGAGCCGGA 6549
Db 2389 TCTGTTCAACACAGCACTTAATTTGTTGCTGAAGAAACCCAGTAAATATATCTGAT 2448
Qy 6550 GTGCAAGAACCGTTCCGTAAATCAATTTGCAAGAAAGCAGAGTGAATGTAAGAA 6609
Db 2449 GTATCAGGCAAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2508
Qy 6610 GCA----AAGATTTTGTGTAAGAAAGCTGATTAACAGAAACGTTAATGTTGTTGTC 6665
Db 2509 GATGGAAGATTAATTAATCTCAATATGCACTGAAGGCCAAAGATTAATCAATCTCATAT 2568
Qy 6724 GCAGTTACCAAGATGAATCAAAACAGAGCAAGATCAAAAT-----TCTAAATTA 6778
Db 2629 GCTATCAACAGGTTGCTCCAAACACACAGAGCTCCAACTACGCGCATTAACATTA 2688
Qy 6779 TGACTGAAACAGT-----TAGATTAATGAGAAATGATTAATTAATCAAGTACT 6831
Db 2689 ACACAGACAGCTGTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2748
Qy 6832 GGAATCGTTCAAGCGAGCTGAATTTCTTGACGCGAGATCTGAGATGATGATGATGATGAT 6891
Db 2749 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2808
Qy 6892 AATTAATTTGCAATTAAGTGAAGAAACAGATTAATGATCAATGATCAATGATCAATGATCAAT 6951

QY	241	AATAATATAGCAAACTCTATATTTTGGGAAAAGAAATAGTACGGGGGTAATAATCTTTTT	300
Db	241	AATAATATAGCAAACTATATTTTGGGAAAAGAAATAGTACGGGGGTAATAATCTTTTT	300
QY	301	AACCTTGCAATGGA AAAATGAA GTAGATGGGATTATCAACGGAAATTCGAGAAAATTA	360
Db	301	AACCTTGCAATGGA AAAATGAA GTAGATGGGATTATCAACGGAAATTCGAGAAAATTA	360
QY	361	ATTGAGAGAAATTTATATTTCTTAAGCTCGAAGGGATGGCAGTAGAGAAAAATGAGTT	420
Db	361	ATTGAGAGAAATTTATATTTCTTAAGCTCGAAGGGATGGCAGTAGAGAAAAATGAGTT	420
QY	421	ATCAATGCTGTTCTTTTCATTCTATTAATCCAAAAAAGATGATTTTAAAGAGCTTTG	480
Db	421	ATCAATGCTGTTCTTTTCATTCTATTAATCCAAAAAAGATGATTTTAAAGAGCTTTG	480
QY	481	GAAGAAGCCAAACATGCTAAGTTTTTAATGGAATCATTCAGATAGATGGA AAAAGTAAA	540
Db	481	GAAGAAGCCAAACATGCTAAGTTTTTAATGGAATCATTCAGATAGATGGA AAAAGTAAA	540
QY	541	ATTCCATTGAATCCGAATGGAAGCATTAACGGTAGAAGAAAAATCAATGCTGTGAAGGC	600
Db	541	ATTCCATTGAATCCGAATGGAAGCATTAACGGTAGAAGAAAAATCAATGCTGTGAAGGC	600
QY	601	ATCGGTTTATATGCGGCGGATTAATGATTAAGTAAGTACTGCAATACTAAGACAGAAAT	660
Db	601	ATCGGTTTATATGCGGCGGATTAATGATTAAGTAAGTACTGCAATACTAAGACAGAAAT	660
QY	661	ACAGATTTTAAAAATTTTAGTCAATTTAGTATCGAATTAATTTCTGGCTGACCGGAGAT	720
Db	661	ACAGATTTTAAAAATTTTAGTCAATTTAGTATCGAATTAATTTCTGGCTGACCGGAGAT	720
QY	721	TTAAAGCTACCAAGACAAAATCTGGAATATTATTTCTTCAGCTCACATGATTTCTCCT	780
Db	721	TTAAAGCTACCAAGACAAAATCTGGAATATTATTTCTTCAGCTCACATGATTTCTCCT	780
QY	781	CAAAAAGCTATGGAAAAAAATTTCAACTGTGGAAGAGAAATAGAATATGTAAAAAGA	840
Db	781	CAAAAAGCTATGGAAAAAAATTTCAACTGTGGAAGAGAAATAGAATATGTAAAAAGA	840
QY	841	AATACCAAGCAAAATATTGAATCTGATGCTGATTGGAGCAGATGGAATATTAATAAT	900
Db	841	AATACCAAGCAAAATATTGAATCTGATGCTGATTGGAGCAGATGGAATATTAATAAT	900
QY	901	AGTGGAAGCTACCAATGGGAGATTTTAAGAAAGAAAGGGGAAAAAGAACTTATPAC	960
Db	901	AGTGGAAGCTACCAATGGGAGATTTTAAGAAAGAAAGGGGAAAAAGAACTTATPAC	960
QY	961	ACTCCTTTAAGTTATTCAGATGTGGAAGCTTCGTAAGATTAATTAAGAAAAATCATA	1020
Db	961	ACTCCTTTAAGTTATTCAGATGTGGAAGCTTCGTAAGATTAATTAAGAAAAATCATA	1020
QY	1021	GGAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTTAAGTTACT	1080
Db	1021	GGAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTTAAGTTACT	1080
QY	1081	AAGCTTGAAAGCACTCTTTTAAGCTTTGTAACAGTTCTATTTCTCCATCAATTTTAAT	1140
Db	1081	AAGCTTGAAAGCACTCTTTTAAGCTTTGTAACAGTTCTATTTCTCCATCAATTTTAAT	1140
QY	1141	GGAATTTTAAGTTTATTTGCAAGTAGTCCAGGTGCTTATTTGGAAAAAGATGCCAAAGTC	1200
Db	1141	GGAATTTTAAGTTTATTTGCAAGTAGTCCAGGTGCTTATTTGGAAAAAGATGCCAAAGTC	1200
QY	1201	GAAGCAACAGAAAGGCAAAATATTCATTTCTTAACAGTGAATGAAGCACTATGGGA	1260
Db	1201	GAAGCAACAGAAAGGCAAAATATTCATTTCTTAACAGTGAATGAAGCACTATGGGA	1260
QY	1261	GAGGCTACTCTTCATTAATAATTAATTCGAATTTATTTGGAGAAAGCCAAATGAAAACTT	1320
Db	1261	GAGGCTACTCTTCATTAATAATTAATTCGAATTTATTTGGAGAAAGCCAAATGAAAACTT	1320

QY	1321	CTCAGTATCGGAGCGGAGTATATTTCTGCAAAAAGTAATTCGATGTAATCTTGAAGGA	1380
Db	1321	CTCAGTATCGGAGCGGAGTATATTTCTGCAAAAAGTAAATTCGATGTAATCTTGAAGGA	1380
QY	1381	GAGTAAATCGAAGGGGAGGAGCAGATATTAATCTCAAAATCTGAAAATPACTATTTGAGTCT	1440
Db	1381	GAGTAAATCGAAGGGGAGGAGCAGATATTAATCTCAAAATCTGAAAATPACTATTTGAGTCT	1440
QY	1441	TCTGTTTCTGTGGAGCGATGAGAGATTCCAAATTAAGTAGCTCTTTCAGTATTTGGTGAACG	1500
Db	1441	TCTGTTTCTGTGGAGCGATGAGAGATTCCAAATTAAGTAGCTCTTTCAGTATTTGGTGAACG	1500
QY	1501	GAAAGAGAAAATTAATCTTCCGTCAGAATTCGTAAGAGCAAAAGTAAGATCGAAAACG	1560
Db	1501	GAAAGAGAAAATTAATCTTCCGTCAGAATTCGTAAGAGCAAAAGTAAGATCGAAAACG	1560
QY	1561	GATCATGTAAATGTGAGAGAGTGAAGCAATTATCCAAATGAGAGTGCCTGTAAAGGTGGA	1620
Db	1561	GATCATGTAAATGTGAGAGAGTGAAGCAATTATTCCAATGAGAGTGCCTGTAAAGGTGGA	1620
QY	1621	TTGGGGGATAGTGGTAAATGGGGTTGTGGCTGCAAAATATTTCTAATCTAATGCTCTCC	1680
Db	1621	TTGGGGGATAGTGGTAAATGGGGTTGTGGCTGCAAAATATTTCTAATCTAATGCTCTCC	1680
QY	1681	CGTATAGATGTAGATGATATCTACATGCCAAGAGCGACTAAATGTGAGAGCTCATTAAC	1740
Db	1681	CGTATAGATGTAGATGATATCTACATGCCAAGAGCGACTAAATGTGAGAGCTCATTAAC	1740
QY	1741	ATTACTAAATTAATGTTCTGCGCAACAGGATCTGAATTTGGGAACTTCCAGCTTATGAAT	1800
Db	1741	ATTACTAAATTAATGTTCTGCGCAACAGGATCTGAATTTGGGAACTTCCAGCTTATGAAT	1800
QY	1801	GATCAGCTTTATGATCAGTGCATCTAAATCAATTTTATGATGCAATTAACACAGCGTTT	1860
Db	1801	GATCAGCTTTATGATCAGTGCATCTAAATCAATTTTATGATGCAATTAACACAGCGTTT	1860
QY	1861	GGAAGAGACAGTGTCAATGAGAAAATTAAGAAATAGCTAACGAATTTATGTGCGGT	1920
Db	1861	GGAAGAGACAGTGTCAATGAGAAAATTAAGAAATAGCTAACGAATTTATGTGCGGT	1920
QY	1921	GTTGTTGCAACCATTAAGCAATTAATTCGTGCTTCTGAGGAATTAAGGAGAGTGGGA	1980
Db	1921	GTTGTTGCAACCATTAAGCAATTAATTCGTGCTTCTGAGGAATTAAGGAGAGTGGGA	1980
QY	1981	AGACTTTCTTCAGAGTGAAGGAGAGTATGTAAAGGCAATTAATGAAGCTCAAAATCTT	2040
Db	1981	AGACTTTCTTCAGAGTGAAGGAGAGTATGTAAAGGCAATTAATGAAGCTCAAAATCTT	2040
QY	2041	CGAGCGACTACGTCAAGTGAAGTGTGGCTGTACGAAAGGAAGAAAAAGAAATTTATTT	2100
Db	2041	CGAGCGACTACGTCAAGTGAAGTGTGGCTGTACGAAAGGAAGAAAAAGAAATTTATTT	2100
QY	2101	GGAATATGACAGATTTTTTATGGAACCTATTAATAATATGCTTCTGTGACAAATTCGCGAT	2160
Db	2101	GGAATATGACAGATTTTTTATGGAACCTATTAATAATATGCTTCTGTGACAAATTCGCGAT	2160
QY	2161	CATGCTGAATTTGGTATCGGAAGAGAAAAATGATATCAACAGTGAATAATTTGAATAT	2220
Db	2161	CATGCTGAATTTGGTATCGGAAGAGAAAAATGATATCAACAGTGAATAATTTGAATAT	2220
QY	2221	AAAAATCCTTCMAAAATGGCAAAAGTCTGTATTTGATTAATTTAGAATTTTAAGAAGACT	2280
Db	2221	AAAAATCCTTCMAAAATGGCAAAAGTCTGTATTTGATTAATTTAGAATTTTAAGAAGACT	2280
QY	2281	TTTGGAAAAAGAAAGAAAAATCTCCAGAAATATGATCGAAAGATATTTGAATCTATTTGA	2340
Db	2281	TTTGGAAAAAGAAAGAAAAATCTCCAGAAATATGATCGAAAGATATTTGAATCTATTTGA	2340
QY	2341	TTATTTGAATGATTTTTCGAAAAATTTGGAATGGAACCGGAGCTTTTATCTAAATGGTGAA	2400
Db	2341	TTATTTGAATGATTTTTCGAAAAATTTGGAATGGAACCGGAGCTTTTATCTAAATGGTGAA	2400
QY	2401	AGAAATGACAAATTAATCTTCCGGAATGGAATTTCAAAAAAGCAAGTCTATTAAGAAATTCGA	2460

Db 2401 AGAATGACAAATTATCTCCGGATGGAACCTTCAAAAACGGAACGTATAGAAATTGCA 2460
Qy 2461 AACTATGTTGAGGGGAAATGAAAAATTAGGAAAAAATTACCGAAGAGATTTAAAGT 2520
Db 2461 AACTATGTTGAGGGGAAATGAAAAATTAGGAAAAAATTACCGAAGAGATTTAAAGT 2520
Qy 2521 TTTTCAGAGAGATGAGTGAATGATTAAGAACTTTGAATTTTACAGAGATAGGAAT 2580
Db 2521 TTTTCAGAGAGATGAGTGAATGATTAAGAACTTTGAATTTTACAGAGATAGGAAT 2580
Qy 2581 TATGCAAAATTTTCAACATTTTACCTCTCCGAGCTAATGAGAAAAGAGATTTCTCT 2640
Db 2581 TATGCAAAATTTTCAACATTTTACCTCTCCGAGCTAATGAGAAAAGAGATTTCTCT 2640
Qy 2641 GTGGAGAGAGCTGTTTCGTGGGTAGAACAGAGAAATTATAGCAAGTATCCGTGGAAA 2700
Db 2641 GTGGAGAGAGCTGTTTCGTGGGTAGAACAGAGAAATTATAGCAAGTATCCGTGGAAA 2700
Qy 2701 GGAGCTAAACTGCTGCAAAAAAAGATTAAATATTAAGCTATCAATTAAGCAGAAACA 2760
Db 2701 GGAGCTAAACTGCTGCAAAAAAAGATTAAATATTAAGCTATCAATTAAGCAGAAACA 2760
Qy 2761 GTGAAATTTAGTTGAAATATTTGCACTTGCAGAGAGCAATACCGAAGTCCAGTCGGA 2820
Db 2761 GTGAAATTTAGTTGAAATATTTGCACTTGCAGAGAGCAATACCGAAGTCCAGTCGGA 2820
Qy 2821 GGAAGATTAATGTTTCAAAAGATCGAAAAATTCACGCTATCTGTAAGACTTAAGAAAAAGCT 2880
Db 2821 GGAAGATTAATGTTTCAAAAGATCGAAAAATTCACGCTATCTGTAAGACTTAAGAAAAAGCT 2880
Qy 2881 GAATTAATCGAGAAAAATTTAATGAGATGCAATTAACAGACTTTTTCATGAGCGGA 2940
Db 2881 GAATTAATCGAGAAAAATTTAATGAGATGCAATTAACAGACTTTTTCATGAGCGGA 2940
Qy 2941 TCTTTTAAATGAGGCTCAGGTGGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGT 3000
Db 2941 TCTTTTAAATGAGGCTCAGGTGGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGT 3000
Qy 3001 ATCAGTAAAGCAAGAGTTTCCATTTGATGACGAAGCATATTTGGAACCTAATAAAAAAT 3060
Db 3001 ATCAGTAAAGCAAGAGTTTCCATTTGATGACGAAGCATATTTGGAACCTAATAAAAAAT 3060
Qy 3061 GCTTTTAAACATTAAGATGATATCTTCTGTTTGAATGCTGCCGTTCAACGGAAATCGGA 3120
Db 3061 GCTTTTAAACATTAAGATGATATCTTCTGTTTGAATGCTGCCGTTCAACGGAAATCGGA 3120
Qy 3121 ACGAAAAATGCGGCGGTGCGGTTGCTGTTGCGGTAAATGATATTTTCAAAACAA 3180
Db 3121 ACGAAAAATGCGGCGGTGCGGTTGCTGTTGCGGTAAATGATATTTTCAAAACAA 3180
Qy 3181 GCTTCCATTTGAAGATTAATGACGAAGCAAAATTAATGATTAAGATTAAGATGATGA 3240
Db 3181 GCTTCCATTTGAAGATTAATGACGAAGCAAAATTAATGATTAAGATTAAGATGATGA 3240
Qy 3241 GTPAACGTAACTGCGGAATCTTTAGAAATGATGCAAAAAAGCCGGAACAATCAACAGT 3300
Db 3241 GTPAACGTAACTGCGGAATCTTTAGAAATGATGCAAAAAAGCCGGAACAATCAACAGT 3300
Qy 3301 ATTTCTGTTGCGGAGAAATTAATTAAGTTGGAAGTAAACCGAGTGAAGAAAAACGAAA 3360
Db 3301 ATTTCTGTTGCGGAGAAATTAATTAAGTTGGAAGTAAACCGAGTGAAGAAAAACGAAA 3360
Qy 3361 TCAGAGAAAGAACAGAGGATTTTTCGCAAAATCGGAACAAAGTGAATCTGTGTA 3420
Db 3361 TCAGAGAAAGAACAGAGGATTTTTCGCAAAATCGGAACAAAGTGAATCTGTGTA 3420
Qy 3421 AATTAATTTGAGATGATGATTTCTTAACGAAAAAATTCAAAATTCATTTCTGAA 3480
Db 3421 AATTAATTTGAGATGATGATTTCTTAACGAAAAAATTCAAAATTCATTTCTGAA 3480
Qy 3481 GAGTAAAAAAGCGGGGAATCTTCTCGAAGTTTCTCATCTCCGATTAAGACCG 3540
Db 3481 GAGTAAAAAAGCGGGGAATCTTCTCGAAGTTTCTCATCTCCGATTAAGACCG 3540

Db 3481 GAGTAAAAAAGCGGGGAATCTTCTCGAAGTTTCTCATCTCCGATTAAGACCG 3540
Qy 3541 TCTTTCACTTTGGAGCTTCTGAAAGTTTCTTCAATTAATTAATAAGAAACATCT 3600
Db 3541 TCTTTCACTTTGGAGCTTCTGAAAGTTTCTTCAATTAATTAATAAGAAACATCT 3600
Qy 3601 GCTGTCTGATGAGTAAAGATTAATTTGAAGGAGCAATTAAGAGTGAAGTACT 3660
Db 3601 GCTGTCTGATGAGTAAAGATTAATTTGAAGGAGCAATTAAGAGTGAAGTACT 3660
Qy 3661 TCTTCTGATTTACTTTTGTGAGACATGGGCGGATCTGCTGCACTTCACTGGAATCAT 3720
Db 3661 TCTTCTGATTTACTTTTGTGAGACATGGGCGGATCTGCTGCACTTCACTGGAATCAT 3720
Qy 3721 ATTGGAATGGAATAATGCAACATCACTGCTGTTAGCTGAGCGGCTGCTTAATAAT 3780
Db 3721 ATTGGAATGGAATAATGCAACATCACTGCTGTTAGCTGAGCGGCTGCTTAATAAT 3780
Qy 3781 ATTCAAAGTAAACCAAGTCTTGGTTAAAAATAGTATATTGAAATGCAATTAATTT 3840
Db 3781 ATTCAAAGTAAACCAAGTCTTGGTTAAAAATAGTATATTGAAATGCAATTAATTT 3840
Qy 3841 AAAGTAAATGCTTTGAGTGAAGAACTCAAGTACACAGAGACAGTTGGAACAGT 3900
Db 3841 AAAGTAAATGCTTTGAGTGAAGAACTCAAGTACACAGAGACAGTTGGAACAGT 3900
Qy 3901 AAAGAAATGAGGACAGAAAAAAGTTATCATTTGGGAATCTTCTTATCACTTA 3960
Db 3901 AAAGAAATGAGGACAGAAAAAAGTTATCATTTGGGAATCTTCTTATCACTTA 3960
Qy 3961 GTGAACAATGAAAGTTTCTGCAAAATCAGAAAAATTAACATGACAGAGAACTGA 4020
Db 3961 GTGAACAATGAAAGTTTCTGCAAAATCAGAAAAATTAACATGACAGAGAACTGA 4020
Qy 4021 CAAAAAATGAGATGTTGATGTCATCTGCTTAATCAAGCGGACACCAAGTGAAGCTTTA 4080
Db 4021 CAAAAAATGAGATGTTGATGTCATCTGCTTAATCAAGCGGACACCAAGTGAAGCTTTA 4080
Qy 4081 AATTTTCAAGCTGGAAGATGCAATGGAATGGAAGGCTTATGAGTGAAGTGAAGTGA 4140
Db 4081 AATTTTCAAGCTGGAAGATGCAATGGAATGGAAGGCTTATGAGTGAAGTGAAGTGA 4140
Qy 4141 AACCAAAAGTAAATGCTTCTTATGAGTGGAGATATATTAAGTTAATGAGCGGAC 4200
Db 4141 AACCAAAAGTAAATGCTTCTTATGAGTGGAGATATATTAAGTTAATGAGCGGAC 4200
Qy 4201 GCAAAAAGCTCTTTTGAACAACAATGATGATGCTGCTGAGTGAACGCGGAGGACAA 4260
Db 4201 GCAAAAAGCTCTTTTGAACAACAATGATGATGCTGCTGAGTGAACGCGGAGGACAA 4260
Qy 4261 AGTTCTGAGCGGAGTTAGGAATTAATCAAGGGGCTTTCTGTCAATTAAGATGCAAT 4320
Db 4261 AGTTCTGAGCGGAGTTAGGAATTAATCAAGGGGCTTTCTGTCAATTAAGATGCAAT 4320
Qy 4321 GACGTGAAGCTAGCGTTGATTAATCTTCAATGAGAGAGCTAATGAATCAATGCAAT 4380
Db 4321 GACGTGAAGCTAGCGTTGATTAATCTTCAATGAGAGAGCTAATGAATCAATGCAAT 4380
Qy 4381 GCCAAAGATGTCAAAAGAAATTTCTGATCTAGCAAAAGAAATTAAGGCTTTACTAA 4440
Db 4381 GCCAAAGATGTCAAAAGAAATTTCTGATCTAGCAAAAGAAATTAAGGCTTTACTAA 4440
Qy 4441 AAAAGTAAAAAATTTTAAAGATGCTGATTTAATACGATCGGAATGCTTAATAGC 4500
Db 4441 AAAAGTAAAAAATTTTAAAGATGCTGATTTAATACGATCGGAATGCTTAATAGC 4500
Qy 4501 AAGGAACAATGAAAAAGCAAAAGAAAAAAGAGAGCGGTCAATTTGAATGCTGCTTTA 4560
Db 4501 AAGGAACAATGAAAAAGCAAAAGAAAAAAGAGAGCGGTCAATTTGAATGCTGCTTTA 4560
Qy 4561 TCGGTTGCTGAACCGATTAATTCGCTGAGAGATGATTTGCAATCTGTTTAA 4620
Db 4561 TCGGTTGCTGAACCGATTAATTCGCTGAGAGATGATTTGCAATCTGTTTAA 4620

Qy	4621	AATAAATTTTAAAGCAGATTGAGTGGAAAGCAATAAGAAAGCCGAGAGAGATPAAATTCAT	4680
Db	4621	AATAAATTTTAAAGCAGATTGAGTGGAAAGCAATAAGAAAGCCGAGAGAGATPAAATTCAT	4680
Qy	4681	GCAGAAACATGTAAATGTGGAGCAAAATCATCTACTGTGTTGTGAAATGGCGGCTTCGGA	4740
Db	4681	GCAGAAACATGTAAATGTGGAGCAAAATCATCTACTGTGTTGTGAAATGGCGGCTTCGGA	4740
Qy	4741	CTTGCTATCAGCAAAAGATGCTTTTTCAGAAATGGGATCTGGAGTCGCAAGACTTATCA	4800
Db	4741	CTTGCTATCAGCAAAAGATGCTTTTTCAGAAATGGGATCTGGAGTCGCAAGACTTATCA	4800
Qy	4801	AATGACACGATTTGCAAAAGTGGATTAAGAGAAATTTCTGCTGATTTCTTAAATGTGAAC	4860
Db	4801	AATGACACGATTTGCAAAAGTGGATTAAGAGAAATTTCTGCTGATTTCTTAAATGTGAAC	4860
Qy	4861	GCAAAATTAATTCATCTTGCGGGTGAATGTGGCGGAAACCATTGCCGGTCTCTTCTACG	4920
Db	4861	GCAAAATTAATTCATCTTGCGGGTGAATGTGGCGGAAACCATTGCCGGTCTCTTCTACG	4920
Qy	4921	GCGGTAGAGAGCTGCTTTTGGCAATATACTCTTCAATPAAAAACCTGCTTTGATTACA	4980
Db	4921	GCGGTAGAGAGCTGCTTTTGGCAATATACTCTTCAATPAAAAACCTGCTTTGATTACA	4980
Qy	4981	GGAACGAAGGTAAATCCTTTTATGTGAAAGAAATACAAAGTCATATGTACAAAGCTTGAT	5040
Db	4981	GGAACGAAGGTAAATCCTTTTATGTGAAAGAAATACAAAGTCATATGTACAAAGCTTGAT	5040
Qy	5041	GATTCATATTAACAACGTTTCTGCTGAGAGCGCTGCAGATATTAAACAGGCTGGAATC	5100
Db	5041	GATTCATATTAACAACGTTTCTGCTGAGAGCGCTGCAGATATTAAACAGGCTGGAATC	5100
Qy	5101	GGAAGAAATGTATCTGTCAATGTGTGTTCTGATGAAACGAAAGCTTTAGTTAGTATCT	5160
Db	5101	GGAAGAAATGTATCTGTCAATGTGTGTTCTGATGAAACGAAAGCTTTAGTTAGTATCT	5160
Qy	5161	GAGTTTGAAGAGTAAAGTTCTTTCAATGTAGATGCAAAAGATCAAAAAACAATAATACA	5220
Db	5161	GAGTTTGAAGAGTAAAGTTCTTTCAATGTAGATGCAAAAGATCAAAAAACAATAATACA	5220
Qy	5221	ATTGCGGAAATGCAAAATGAGAAAGACGCGTGGAGTTGGAGCAACAGTTGCTCATACA	5280
Db	5221	ATTGCGGAAATGCAAAATGAGAAAGACGCGTGGAGTTGGAGCAACAGTTGCTCATACA	5280
Qy	5281	AATATTTGAAAAACAATCACTTATAGCTATTGTAAAAACAGTAAATTTACAACGGCGAAT	5340
Db	5281	AATATTTGAAAAACAATCACTTATAGCTATTGTAAAAACAGTAAATTTACAACGGCGAAT	5340
Qy	5341	GATCAAGATAGAAAAAATATCATGTGACTGCAAAAAGATTACTATGACCAATACTATA	5400
Db	5341	GATCAAGATAGAAAAAATATCATGTGACTGCAAAAAGATTACTATGACCAATACTATA	5400
Qy	5401	GCAGTCGGAATGTTGAGAGCAAAAGAGACCTCTGTGCAAGAGCTTTCGAAGTACTACC	5460
Db	5401	GCAGTCGGAATGTTGAGAGCAAAAGAGACCTCTGTGCAAGAGCTTTCGAAGTACTACC	5460
Qy	5461	TTGAATTAAGACAGTTTCTTCTCATGTGTATCAAACTGATATTGACAAAGATTTAGAGAA	5520
Db	5461	TTGAATTAAGACAGTTTCTTCTCATGTGTATCAAACTGATATTGACAAAGATTTAGAGAA	5520
Qy	5521	GAAAAATTAATGGAATTAAGAAAAAGCAAAATGTAAATGTTCTTAGCTGAAATACGAGTCAA	5580
Db	5521	GAAAAATTAATGGAATTAAGAAAAAGCAAAATGTAAATGTTCTTAGCTGAAATACGAGTCAA	5580
Qy	5581	GTGTGCACAAATGCGACAGTGCTTTCCGAGACAGTGGACAAAGCTGCAGTATGGAAGCTGGA	5640
Db	5581	GTGTGCACAAATGCGACAGTGCTTTCCGAGACAGTGGACAAAGCTGCAGTATGGAAGCTGGA	5640
Qy	5641	GTAGCAGTTTAATTAATTAACAACAAATACTCTGACATATPAAAAAATAGTACTCAAAAT	5700
Db	5641	GTAGCAGTTTAATTAATTAACAACAAATACTCTGACATATPAAAAAATAGTACTCAAAAT	5700

Db 6781 ACTCGAAACAATTGATGTAATAGCAGAAATGATTAATTCAGATCTGGATCCGT 6840
Qy 6841 TCAGCCGAGAGCTGGAATTCCTTCAGCCGAGATCTGAGTGGTTCCTCAATATAT 6900
Db 6841 TCAGCCGAGAGCTGGAATTCCTTCAGCCGAGATCTGAGTGGTTCCTCAATATAT 6900
Qy 6901 GCMAATTAAGGTGAAACAGATATCGATCTACTCTTCACTCTTCAATGTAAT 6960
Db 6901 GCMAATTAAGGTGAAACAGATATCGATCTACTCTTCACTCTTCAATGTAAT 6960
Qy 6961 GTPAAAGCTTAAATPAAATTCGAAATTCCTTGAACAGCCGCTGAGAGCCGAGTCT 7020
Db 6961 GTPAAAGCTTAAATPAAATTCGAAATTCCTTGAACAGCCGCTGAGAGCCGAGTCT 7020
Qy 7021 GCAGCAGTACCGAGAGTGGTTCCTGTTAACTATTAATAGTTCGTGATGCTGAGT 7080
Db 7021 GCAGCAGTACCGAGAGTGGTTCCTGTTAACTATTAATAGTTCGTGATGCTGAGT 7080
Qy 7081 CACAACTACTGATTTGACTTCCTGTAAGGAAAAAGTAAATGTAACGGCAAAAGGAA 7140
Db 7081 CACAACTACTGATTTGACTTCCTGTAAGGAAAAAGTAAATGTAACGGCAAAAGGAA 7140
Qy 7141 AAAAATATTAGCAAAACAGCAGCAAAATGAGGAATCGAGAGCAGCAATCGAGCCAT 7200
Db 7141 AAAAATATTAGCAAAACAGCAGCAAAATGAGGAATCGAGAGCAGCAATCGAGCCAT 7200
Qy 7201 GTCTTGTAATTAATTTTGGAAACAGCTGTAAGATAGAAAAATTCGAAAGAAAAAG 7260
Db 7201 GTCTTGTAATTAATTTTGGAAACAGCTGTAAGATAGAAAAATTCGAAAGAAAAAG 7260
Qy 7261 ACAGAAATTTTAAAACTTTAGACGAGTTAACAAAGAAACAGATAAAAAGTAAATGAT 7320
Db 7261 ACAGAAATTTTAAAACTTTAGACGAGTTAACAAAGAAACAGATAAAAAGTAAATGAT 7320
Qy 7321 GGTAGAAAAAATCTTACATCAGAGGATTTCTACAGAAATGACTCTGTAAGAAAG 7380
Db 7321 GGTAGAAAAAATCTTACATCAGAGGATTTCTACAGAAATGACTCTGTAAGAAAG 7380
Qy 7381 GATAGAGAGATTAATCAGGAGAGAAATTAAGCCATTTGTGAAGCTTCGATATAT 7440
Db 7381 GATAGAGAGATTAATCAGGAGAGAAATTAAGCCATTTGTGAAGCTTCGATATAT 7440
Qy 7441 GGAATAATGTAATTAACACAGAGACAGATTAATCACTCTGATGATGTTG 7500
Db 7441 GGAATAATGTAATTAACACAGAGACAGATTAATCACTCTGATGATGTTG 7500
Qy 7501 GGAATGCGAGGCTTCTCCGATCAGAAACAGTGGCAGTTAACAAATTTTAAAGAAAT 7560
Db 7501 GGAATGCGAGGCTTCTCCGATCAGAAACAGTGGCAGTTAACAAATTTTAAAGAAAT 7560
Qy 7561 TCCGAGTACTGTTGAAAAATCTTTTGTGAAGACAGTGAATAATGTTAGATCG 7620
Db 7561 TCCGAGTACTGTTGAAAAATCTTTTGTGAAGACAGTGAATAATGTTAGATCG 7620
Qy 7621 GATATTACAGGAAATGTTGTTTAAACAGCATATCAAGTCTCTAGAGAGCATGGGAATA 7680
Db 7621 GATATTACAGGAAATGTTGTTTAAACAGCATATCAAGTCTCTAGAGAGCATGGGAATA 7680
Qy 7681 GAGAGCTGCTATGCAAGATTAATTCATAGGAAGATCAATATCATATTAATAATTC 7740
Db 7681 GAGAGCTGCTATGCAAGATTAATTCATAGGAAGATCAATATCATATTAATAATTC 7740
Qy 7741 AAGCTATTAGAAAAATATTTGATGTTATTTAAAGATTAATCGAATTTGAGCGGAA 7800
Db 7741 AAGCTATTAGAAAAATATTTGATGTTATTTAAAGATTAATCGAATTTGAGCGGAA 7800
Qy 7801 GCAAAAGATTAACCGTAGAGCGGTAGCTGCGAGAGCCATATCTTAAAGCAAGAAAT 7860
Db 7801 GCAAAAGATTAACCGTAGAGCGGTAGCTGCGAGAGCCATATCTTAAAGCAAGAAAT 7860
Qy 7861 GAATGAATTCAGAGGTGAAATTTGAGAGAGTATTTCAATGAAGAAATAGAGTACT 7920
Db 7861 GAATGAATTCAGAGGTGAAATTTGAGAGAGTATTTCAATGAAGAAATAGAGTACT 7920

Db 7861 GAATGAATTCAGAGGTGAAATTTGAGAGAGTATTTCAATGAAGAAATAGAGTACT 7920
Qy 7921 AGCCCTTAAAGAAATTTGAAAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGT 7980
Db 7921 AGCCCTTAAAGAAATTTGAAAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGT 7980
Qy 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTACGAGGCGCAGAAATTAATTCGAA 8040
Db 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTACGAGGCGCAGAAATTAATTCGAA 8040
Qy 8041 GCAAAAGATGCGGAGCTTATTTGAAAGTATGACAAATCCGAAAGAAATTTT 8100
Db 8041 GCAAAAGATGCGGAGCTTATTTGAAAGTATGACAAATCCGAAAGAAATTTT 8100
Qy 8101 CATGCAATTAATGCAATATGGAAGCAACATTAATGAAGTAAACGAGCTTCTAA 8160
Db 8101 CATGCAATTAATGCAATATGGAAGCAACATTAATGAAGTAAACGAGCTTCTAA 8160
Qy 8161 GCAGTAAAGGTCTGTATTGGAGAGTGGAGTCAACAGGCAAGAGTACTGCTGCA 8220
Db 8161 GCAGTAAAGGTCTGTATTGGAGAGTGGAGTCAACAGGCAAGAGTACTGCTGCA 8220
Qy 8221 GGTAAACTATGTAAGAAAGTTGAGAGAAATTTGTCAGAACAAATCGATTGAATGCA 8280
Db 8221 GGTAAACTATGTAAGAAAGTTGAGAGAAATTTGTCAGAACAAATCGATTGAATGCA 8280
Qy 8281 ATTTTAAAGTGAAGGTTTGAATGAATTAATGTAATGTTCTGATGATCA 8340
Db 8281 ATTTTAAAGTGAAGGTTTGAATGAATTAATGTAATGTTCTGATGATCA 8340
Qy 8341 GGAATGAGAGAGATTTGCGAGCAGAGTGAATCTTCAACAGCAAAAGTAACT 8400
Db 8341 GGAATGAGAGAGATTTGCGAGCAGAGTGAATCTTCAACAGCAAAAGTAACT 8400
Qy 8401 GAATCCGATGTTGTTTACGAAAGCAAGTTATGAAATATGATTACACAAAAATAT 8460
Db 8401 GAATCCGATGTTGTTTACGAAAGCAAGTTATGAAATATGATTACACAAAAATAT 8460
Qy 8461 ATTTGAAAGTGAATGCTTCTGTTTAAATGATTAACAAAGAAAGGCGAATTAAGATCT 8520
Db 8461 ATTTGAAAGTGAATGCTTCTGTTTAAATGATTAACAAAGAAAGGCGAATTAAGATCT 8520
Qy 8521 TTAGCGGTAGCCGAGTGTGATGCAACAGGAAACAAAGCAATTTACAGATCAAAAG 8580
Db 8521 TTAGCGGTAGCCGAGTGTGATGCAACAGGAAACAAAGCAATTTACAGATCAAAAG 8580
Qy 8581 TTAACTTCAACATGTAATGAGAGAAACGATATCTCACTTCGTGCAAAAGCTTTGCT 8640
Db 8581 TTAACTTCAACATGTAATGAGAGAAACGATATCTCACTTCGTGCAAAAGCTTTGCT 8640
Qy 8641 AAAAATGAATAATTAAGAAATGTAAGAACTGAGAGAGCTTACTCGAAGCGGAAAC 8700
Db 8641 AAAAATGAATAATTAAGAAATGTAAGAACTGAGAGAGCTTACTCGAAGCGGAAAC 8700
Qy 8701 GAGCGGTGAAATTAATCAAGAGTACTACAGAGAGCAATGTTGTCAGAAATTTGGGAA 8760
Db 8701 GAGCGGTGAAATTAATCAAGAGTACTACAGAGAGCAATGTTGTCAGAAATTTGGGAA 8760
Qy 8761 ATTTGAGATTAATTAAGAAACGATGCAAGAGTAAATGATGTAAGATCAACGAGAC 8820
Db 8761 ATTTGAGATTAATTAAGAAACGATGCAAGAGTAAATGATGTAAGATCAACGAGAC 8820
Qy 8821 GGAACCAAGAGGCTTGTTCGAAAGAAATGTAATGTAATGTAATGTAATGTAATGTAAT 8880
Db 8821 GGAACCAAGAGGCTTGTTCGAAAGAAATGTAATGTAATGTAATGTAATGTAATGTAAT 8880
Qy 8881 GAAACCAATCAATCAATGTAAGTAAGCAGAAATTTGTAAGCAGAAAGTGAATGTA 8940
Db 8881 GAAACCAATCAATCAATGTAAGTAAGCAGAAATTTGTAAGCAGAAAGTGAATGTA 8940
Qy 8941 GATGCTTTGAATGAATTTGATGATCTACAAAGAAAGTGGTGGTATGATGTAAT 9000
Db 8941 GATGCTTTGAATGAATTTGATGATCTACAAAGAAAGTGGTGGTATGATGTAAT 9000

Db 1994 ACTCCTTAAGTTTATCAGATGTGAGAGCTTCCGTAAAGATTAATAAGAAAGATCAATA 2053
Qy 1021 GAAAAGAAATGTTGACATTAACGCTGAGCAAGAAATTTCTATGATGCACTTTACT 1080
Db 2054 GAAAAGAAATGTTGACATTAACGCTGAGCAAGAAATTTCTATGATGCACTTTACT 2113
Qy 1081 AAGCTTGAAGACACTTTTATGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAAT 1140
Db 2114 AAGCTTGAAGACACTTTTATGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAAT 2173
Qy 1141 GGAATTTTATGCTTTATTTGACAAATAGTCCAGTGTCTTATTTGAAAAGATCCAAAGTC 1200
Db 2174 GGAATTTTATGCTTTATTTGACAAATAGTCCAGTGTCTTATTTGAAAAGATCCAAAGTC 2233
Qy 1201 GAAGCAACAGAGAGAAAGCAAAATATTCATTCTTACAGTGTGATGAGCACTATGGA 1260
Db 2234 GAAGCAACAGAGAGAAAGCAAAATATTCATTCTTACAGTGTGATGAGCACTATGGA 2293
Qy 1261 GGAAGTACTTCTCCATTAATAATTTACCAATTTATTTGGAGAAAGCCAAATGAAAACCT 1320
Db 2294 GGAAGTACTTCTCCATTAATAATTTACCAATTTATTTGGAGAAAGCCAAATGAAAACCT 2353
Qy 1321 CTCAGATCGAGCGGATATATTTCTGCAAAAAGTAAATTCATGATGTAATTTGAAGA 1380
Db 2354 CTCAGATCGAGCGGATATATTTCTGCAAAAAGTAAATTCATGATGTAATTTGAAGA 2413
Qy 1381 GAAGTAAATTCAGAGGAGAGAGAGATATTTACTTCAAAATCTGAAAATATCTATGTGCT 1440
Db 2414 GAAGTAAATTCAGAGGAGAGAGAGATATTTACTTCAAAATCTGAAAATATCTATGTGCT 2473
Qy 1441 TCTGTTTCTGTTGGAACGATGAGAGATTCCAATTAAGTACTTTCAATATTTGGAGAG 1500
Db 2474 TCTGTTTCTGTTGGAACGATGAGAGATTCCAATTAAGTACTTTCAATATTTGGAGAG 2533
Qy 1501 GAAGGAGAAATTAATCTTCCGTCAGATGTGCTAAAGAGCAAAAAGTAAATCAGAAAG 1560
Db 2534 GAAGGAGAAATTAATCTTCCGTCAGATGTGCTAAAGAGCAAAAAGTAAATCAGAAAG 2593
Qy 1561 GATGATGTAATGTGAGAAAGTGAAGCGATTAATTCATTGAGCTGCTGTAAGGTGGA 1620
Db 2594 GATGATGTAATGTGAGAAAGTGAAGCGATTAATTCATTGAGCTGCTGTAAGGTGGA 2653
Qy 1621 TTGGGGGATAGTGTAAATGGGGTTGGGCTGCAAAATTTCTTAATCTTAATGCTTCTCC 1680
Db 2654 TTGGGGGATAGTGTAAATGGGGTTGGGCTGCAAAATTTCTTAATCTTAATGCTTCTCC 2713
Qy 1681 CGATAGATGTAATGATATCTTACATGCCAAGAGCGACTAAATGTGAGGCTCATTAAC 1740
Db 2714 CGATAGATGTAATGATATCTTACATGCCAAGAGCGACTAAATGTGAGGCTCATTAAC 2773
Qy 1741 ATTACTTAATAATAGTCTTCTGCAAAAGAGATCTGATTTGGAACTTCCAAATTTATGAAT 1800
Db 2774 ATTACTTAATAATAGTCTTCTGCAAAAGAGATCTGATTTGGAACTTCCAAATTTATGAAT 2833
Qy 1801 GATCAGCTTTATGATCAGTCACTTAATCAATTTAGATGCAATTAAGCGGCTT 1860
Db 2834 GATCAGCTTTATGATCAGTCACTTAATCAATTTAGATGCAATTAAGCGGCTT 2893
Qy 1861 GAGAGAGACAGTGTCAATGAGAAATTAAGATTAAGTAAAGCACTTATTTAGTGTGCT 1920
Db 2894 GAGAGAGACAGTGTCAATGAGAAATTAAGATTAAGTAAAGCACTTATTTAGTGTGCT 2953
Qy 1921 GTGTCTGCAACCAATAGCAAAATCAATTAATTTCTGCTTCTGTGCAATAGAGAGAGTGA 1980
Db 2954 GTGTCTGCAACCAATAGCAAAATCAATTAATTTCTGCTTCTGTGCAATAGAGAGAGTGA 3013
Qy 1981 AGACTTTCTTCAAGAGTGTGAAGGAGATTAATGTAAGGCAATTAATGAAGCTCAAAATCTT 2040
Db 3014 AGACTTTCTTCAAGAGTGTGAAGGAGATTAATGTAAGGCAATTAATGAAGCTCAAAATCTT 3073
Qy 2041 CGAGCGACTAGTCAAGTGTGAAGTGTGGCTGTACGAAAGGAAAGAAAGAACTTAT 2100
Db 3074 CGAGCGACTAGTCAAGTGTGAAGTGTGGCTGTACGAAAGGAAAGAAAGAACTTAT 3133

Qy 2101 GGAATGCAAGAGTTTATTTATGAAAACTATTAATAATAGCTTCTGTGACAAATGCCAT 2160
Db 3134 GGAATGCAAGAGTTTATTTATGAAAACTATTAATAATAGCTTCTGTGACAAATGCCAT 3193
Qy 2161 CATGCTGAATTTGATGCAAGAGAAATTTGATTCAAAGTGAATTAATTAATTAAT 2220
Db 3194 CATGCTGAATTTGATGCAAGAGAAATTTGATTCAAAGTGAATTAATTAATTAAT 3253
Qy 2221 AAAAATCCTTCAAAAATGGCAAAAGCTCTTATTTGTAATTAAGAACTTTAAAGAGCT 2280
Db 3254 AAAAATCCTTCAAAAATGGCAAAAGCTCTTATTTGTAATTAAGAACTTTAAAGAGCT 3313
Qy 2281 TTTGGAAAAGAAAGAAATCCAGAAATATGATCCGAAAGATATTAATCTATGAAAAA 2340
Db 3314 TTTGGAAAAGAAAGAAATCCAGAAATATGATCCGAAAGATATTAATCTATGAAAAA 3373
Qy 2341 TTATTAATGCAATTTTCAAGAAAATTTGATGAAAAACCGAGCTTTTACTAATAGTGA 2400
Db 3374 TTATTAATGCAATTTTCAAGAAAATTTGATGAAAAACCGAGCTTTTACTAATAGTGA 3433
Qy 2401 AGAATGCAATTTATTTCTCCGATGGAATCTTCAAAAACAGGAATGCTATAGAAATGGA 2460
Db 3434 AGAATGCAATTTATTTCTCCGATGGAATCTTCAAAAACAGGAATGCTATAGAAATGGA 3493
Qy 2461 AACTATGTTCAAGGAGAAATGAAAAATTTAGAGAAAAATTCAGAAAAGATTTAAAGCT 2520
Db 3494 AACTATGTTCAAGGAGAAATGAAAAATTTAGAGAAAAATTCAGAAAAGATTTAAAGCT 3553
Qy 2521 TTTTCAGAAAGATTTAGTGAATGTAATTAAGAAATTTGAATTTTACAGAGATAGAAAT 2580
Db 3554 TTTTCAGAAAGATTTAGTGAATGTAATTAAGAAATTTGAATTTTACAGAGATAGAAAT 3613
Qy 2581 TATGCAATTTTCAACTTTTACCTTCTCCGAGCTAATGAGAAAGAGATGTTCTTCT 2640
Db 3614 TATGCAATTTTCAACTTTTACCTTCTCCGAGCTAATGAGAAAGAGATGTTCTTCT 3673
Qy 2641 GTGGAGAGAGCTTTTCTGAGGTAGAACAGAGAAATTAAGCAAGATCCGTTGAAAA 2700
Db 3674 GTGGAGAGAGCTTTTCTGAGGTAGAACAGAGAAATTAAGCAAGATCCGTTGAAAA 3733
Qy 2701 GGAGCTTAACCTTCTGCAAAAAGAAATTTAAATTAATTAAGCTATCAATTAAGCAGAAAC 2760
Db 3734 GGAGCTTAACCTTCTGCAAAAAGAAATTTAAATTAATTAAGCTATCAATTAAGCAGAAAC 3793
Qy 2761 GTGAATTTAGTTGAAATATTTGACCTTGGAGAGAGCAATCCGGAATGCAATCCGA 2820
Db 3794 GTGAATTTAGTTGAAATATTTGACCTTGGAGAGAGCAATCCGGAATGCAATCCGA 3853
Qy 2821 GGAAGTTAAATGTTCAAAAGATCGAAAAATTCAGCTATGTAAGCTTAAGAAAAAGCT 2880
Db 3854 GGAAGTTAAATGTTCAAAAGATCGAAAAATTCAGCTATGTAAGCTTAAGAAAAAGCT 3913
Qy 2881 GAATTAATCAGAGAAAAATTTAATGAGATGATGGAACAGACTTTTTCATGTAGCGGA 2940
Db 3914 GAATTAATCAGAGAAAAATTTAATGAGATGATGGAACAGACTTTTTCATGTAGCGGA 3973
Qy 2941 TCTTTTAAATGTGGCTCAGGTGGAATGCAATCAATGAAATGGAAGTTATAGTGAAGT 3000
Db 3974 TCTTTTAAATGTGGCTCAGGTGGAATGCAATCAATGAAATGGAAGTTATAGTGAAGT 4033
Qy 3001 ATCAGTAAGCAAGAGTTTCCATGATGAGCAAGCAATTTTGAAGCACTTAATAAAAAT 3060
Db 4034 ATCAGTAAGCAAGAGTTTCCATGATGAGCAAGCAATTTTGAAGCACTTAATAAAAAT 4093
Qy 3061 GCTTTAAACAGTAAGATATCTTCTGTTTGAATGCTGCGGTTCAAGCGGAATCCGA 3120
Db 4094 GCTTTAAACAGTAAGATATCTTCTGTTTGAATGCTGCGGTTCAAGCGGAATCCGA 4153
Qy 3121 ACGAAAAATGCGGCTCGGGTTGCTGTGGCGTTAAATGATTAATTTCAACAAA 3180
Db 4154 ACGAAAAATGCGGCTCGGGTTGCTGTGGCGTTAAATGATTAATTTCAACAAA 4213

OY 3181 GCTTCATTGAAGATATGACGAGGACAAAGTAAATATGATTAAGATATGAATGA 3240
| | | | |
DB 4214 GCTTCATTGAAGATATGACGAGGACAAAGTAAATATGATTAAGATATGAATGA 4273
| | | | |
OY 3241 GTPAACGTATCTGCGGAATCTTTAAGAGTATGCAAAAAAGACCGGAACTATCAACGT 3300
| | | | |
DB 4274 GTPAACGTATCTGCGGAATCTTTAAGAGTATGCAAAAAAGACCGGAACTATCAACGT 4333
| | | | |
OY 3301 ATTTCTGTTGCGGAGGAATTAATAAGTTGGAAGTAAACGAGTGAAGAAAAACCGAAA 3360
| | | | |
DB 4334 ATTTCTGTTGCGGAGGAATTAATAAGTTGGAAGTAAACGAGTGAAGAAAAACCGAAA 4393
| | | | |
OY 3361 TCAGAAAGAAAGACGAGAGGATTTTTTGGCAAAATCGAAACAAAGTGACTCTGTAAA 3420
| | | | |
DB 4394 TCAGAAAGAAAGACGAGAGGATTTTTTGGCAAAATCGAAACAAAGTGACTCTGTAAA 4453
| | | | |
OY 3421 AATTAATAATACGATATGATATGATTCATTAACAGAAAAATTAACAATTAATCTGAA 3480
| | | | |
DB 4454 AATTAATAATACGATATGATATGATTCATTAACAGAAAAATTAACAATTAATCTGAA 4513
| | | | |
OY 3481 GGAGTAAAAAAGACGCGGGAATCTTCTTGAAAGTTTCTCATACTCCGATAAAGACCG 3540
| | | | |
DB 4514 GGAGTAAAAAAGACGCGGGAATCTTCTTGAAAGTTTCTCATACTCCGATAAAGACCG 4573
| | | | |
OY 3541 TCTTTCAGTTTGGAGCTTCTGGAAGTGTTCCTTCAATATATTAATAAAGAAACATCT 3600
| | | | |
DB 4574 TCTTTCAGTTTGGAGCTTCTGGAAGTGTTCCTTCAATATATTAATAAAGAAACATCT 4633
| | | | |
OY 3601 GCTGTCGTAGATGAGTAAAGATTAATTTGAAGGAGCAATTAAGTAAAGTGAATCT 3660
| | | | |
DB 4634 GCTGTCGTAGATGAGTAAAGATTAATTTGAAGGAGCAATTAAGTAAAGTGAATCT 4693
| | | | |
OY 3661 TCTTCTGATTTCTACTTTTGTGAGCATGCGGCGATCTGTCGACTTCAGTGAATCAT 3720
| | | | |
DB 4694 TCTTCTGATTTCTACTTTTGTGAGCATGCGGCGATCTGTCGACTTCAGTGAATCAT 4753
| | | | |
OY 3721 ATTTGAAGTGAAGATATGCAACATCACTGCTGTTTATGCTGAGCGGCTCTGTAAATAT 3780
| | | | |
DB 4754 ATTTGAAGTGAAGATATGCAACATCACTGCTGTTTATGCTGAGCGGCTCTGTAAATAT 4813
| | | | |
OY 3781 ATTTCAAGTGAAGATATGCAACATCACTGCTGTTTATGCTGAGCGGCTCTGTAAATAT 3840
| | | | |
DB 4814 ATTTCAAGTGAAGATATGCAACATCACTGCTGTTTATGCTGAGCGGCTCTGTAAATAT 4873
| | | | |
OY 3841 AAAGTAAATGCTTTGAGTGAAGAACTCAAGTAGCAGCAGAGAGTGTGAAGCAAT 3900
| | | | |
DB 4874 AAAGTAAATGCTTTGAGTGAAGAACTCAAGTAGCAGCAGAGAGTGTGAAGCAAT 4933
| | | | |
OY 3901 AAAGAAAGTGAAGGACAGAGAAAAATTAATCTAATTTGGAACCTTCTGTCAACTTA 3960
| | | | |
DB 4934 AAAGAAAGTGAAGGACAGAGAAAAATTAATCTAATTTGGAACCTTCTGTCAACTTA 4993
| | | | |
OY 3961 GTGAACATATGAAGTTTCTGCAAAATGCAAAAAATTAATCAAGTAGCAGAGAACTGAAAGC 4020
| | | | |
DB 4994 GTGAACATATGAAGTTTCTGCAAAATGCAAAAAATTAATCAAGTAGCAGAGAACTGAAAGC 5053
| | | | |
OY 4021 CAAAAAATGAGATGTTGATGCTCACTGTTATCAAGCGGACACCAAGTAGCAGAGAGCTTTA 4080
| | | | |
DB 5054 CAAAAAATGAGATGTTGATGCTCACTGTTATCAAGCGGACACCAAGTAGCAGAGAGCTTTA 5113
| | | | |
OY 4081 AATTTACAAAGTGAAGATCAATGGAACGTATAGGGGCTACTGTGACTGTGCTCAATTA 4140
| | | | |
DB 5114 AATTTACAAAGTGAAGATCAATGGAACGTATAGGGGCTACTGTGACTGTGCTCAATTA 5173
| | | | |
OY 4141 AACCAAAAGTAAATGCTTCTATATAGTGTGGAAGATATACCTTAATGAGAGGAC 4200
| | | | |
DB 5174 AACCAAAAGTAAATGCTTCTATATAGTGTGGAAGATATACCTTAATGAGAGGAC 5233
| | | | |
OY 4201 GCAAAAGCTCTTTTACCAACCACTCAAGTGACTGTGAGTGAAGCAGGAGGAGCAAT 4260
| | | | |
DB 5234 GCAAAAGCTCTTTTACCAACCACTCAAGTGACTGTGAGTGAAGCAGGAGGAGCAAT 5293
| | | | |
OY 4261 AGTTCTGAGCGGAGTTAGAAATTAACAAGGGGCTGTTTCTGTCAATTAAGATTGAACAT 4320
| | | | |

DB 5294 AGTTCTGAGCGGAGTTTGAAGAAATTAACAAGGGGCTGTTTCTGTCAATTAAGATTGACAAAT 5353
| | | | |
OY 4321 GACGTGAAGGCTAGCGTTGATTAATCTTCCATTCGAGAGGCTAATGAATGAATGTCATT 4380
| | | | |
DB 5354 GACGTGAAGGCTAGCGTTGATTAATCTTCCATTCGAGAGGCTAATGAATGAATGTCATT 5413
| | | | |
OY 4381 GCCAAAGATGTCAAGAGAGTTCTGATCTAGCAAAAGATATCAAGCTTTACTAAATGGA 4440
| | | | |
DB 5414 GCCAAAGATGTCAAGAGAGTTCTGATCTAGCAAAAGATATCAAGCTTTACTAAATGGA 5473
| | | | |
OY 4441 AAAGTAAAAAATTTTGAAGATCTGTGTATTAATACGACTGGAATGTTTATTAACG 4500
| | | | |
DB 5474 AAAGTAAAAAATTTTGAAGATCTGTGTATTAATACGACTGGAATGTTTATTAACG 5533
| | | | |
OY 4501 AAAGAACACTAGAAAAAGCAAGAAAAAGAGAGGCGTCAATGTAAATGCTGCTTA 4560
| | | | |
DB 5534 AAAGAACACTAGAAAAAGCAAGAAAAAGAGAGGCGTCAATGTAAATGCTGCTTA 5593
| | | | |
OY 4561 TCGGTTGCTGGAACGGAATTAATCCGCTGAGAGAGTACTATTGCAATTAATGTTAAA 4620
| | | | |
DB 5594 TCGGTTGCTGGAACGGAATTAATCCGCTGAGAGAGTACTATTGCAATTAATGTTAAA 5653
| | | | |
OY 4621 AATTAATTTAAAGCAAAATTTGAGTGAAGCAATTAAGAAAGCCGAGAGATTAATTAAT 4680
| | | | |
DB 5654 AATTAATTTAAAGCAAAATTTGAGTGAAGCAATTAAGAAAGCCGAGAGATTAATTAAT 5713
| | | | |
OY 4681 GCGAAACATGTAAATGTGGAAGCAAAATCATCTACTGTTGTGTAATGCGGCTCTGGA 4740
| | | | |
DB 5714 GCGAAACATGTAAATGTGGAAGCAAAATCATCTACTGTTGTGTAATGCGGCTCTGGA 5773
| | | | |
OY 4741 CTGCTATCAACAAAGATGCTTTTTCAGAAATGGGATCTGAGCATGAGCAAGACTTAACA 4800
| | | | |
DB 5774 CTGCTATCAACAAAGATGCTTTTTCAGAAATGGGATCTGAGCATGAGCAAGACTTAACA 5833
| | | | |
OY 4801 AATGACAGATTTGCAAGAGTGAATTAAGAAAGATTTCTGCTGATTCCTTAATGTGAAC 4860
| | | | |
DB 5834 AATGACAGATTTGCAAGAGTGAATTAAGAAAGATTTCTGCTGATTCCTTAATGTGAAC 5893
| | | | |
OY 4861 GCAAAATTAATTCATTTCTGGGGGTGAATGTTGCGGGAACCATGCGCGGTTCTTTCTACG 4920
| | | | |
DB 5894 GCAAAATTAATTCATTTCTGGGGGTGAATGTTGCGGGAACCATGCGCGGTTCTTTCTACG 5953
| | | | |
OY 4921 GCGGTAGAGGCTGCTTTTTCGCAATTAATCTCTTCAATTAATAAACCCTGCTTTGATTAACA 4980
| | | | |
DB 5954 GCGGTAGAGGCTGCTTTTTCGCAATTAATCTCTTCAATTAATAAACCCTGCTTTGATTAACA 6013
| | | | |
OY 4981 GGAACGAAGTAAATCTTTTATGAGAAAGATTAACAAAGTCAATGTACAAAGCTTTGAAT 5040
| | | | |
DB 6014 GGAACGAAGTAAATCTTTTATGAGAAAGATTAACAAAGTCAATGTACAAAGCTTTGAAT 6073
| | | | |
OY 5041 GATTCTCATTTTAAACAAAGTTTCTGCTGAGAGCGCTGCAAGTATTAACAGGCTGGAATC 5100
| | | | |
DB 6074 GATTCTCATTTTAAACAAAGTTTCTGCTGAGAGCGCTGCAAGTATTAACAGGCTGGAATC 6133
| | | | |
OY 5101 GGAGGAATGATATCTGTCAATCGTGGTTCTGATGAACGGAAGCTTTAGTTAGTATCT 5160
| | | | |
DB 6134 GGAGGAATGATATCTGTCAATCGTGGTTCTGATGAACGGAAGCTTTAGTTAGTATCT 6193
| | | | |
OY 5161 GAGTTGAAGAGTAAAGTTCTTTCAATGATGACAAAGATCAAAAAACATTAATTAACA 5220
| | | | |
DB 6194 GAGTTGAAGAGTAAAGTTCTTTCAATGATGACAAAGATCAAAAAACATTAATTAACA 6253
| | | | |
OY 5221 ATTTGCGGAATGTCAAAATGAGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATCA 5280
| | | | |
DB 6254 ATTTGCGGAATGTCAAAATGAGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATCA 6313
| | | | |
OY 5281 AATATTGAAAAACATCAGTTATAGCTATTGTAAAAAAACAGTAAATTAACAACGCGAAT 5340
| | | | |
DB 6314 AATATTGAAAAACATCAGTTATAGCTATTGTAAAAAAACAGTAAATTAACAACGCGAAT 6373
| | | | |
OY 5341 GATCAAGATTAAGAAAAATATCAATGTGACTGCAAAAGATTAATCTATACCAATTACTATA 5400
| | | | |

Dh 6374 GATCAAGATAGAAAAATATCAATGTGACTGCAAAAAGATTATATGCACTATCTATA 6433
Qy 5401 GCAAGTCGAGTTGGAGAGCAAAAGAGCTCTGGAGCAAGAGCTTTCGCAATGTACTACC 5460
Dh 6434 GCAAGTCGAGTTGGAGAGCAAAAGAGCTCTGGAGCAAGAGCTTTCGCAATGTACTACC 6493
Qy 5461 TTGAATTAAGACAGTTTCTCTCATGTGTGATCAAACTGATTAATGACAAAAGATTAGAGAA 5520
Dh 6494 TTGAATTAAGACAGTTTCTCTCATGTGTGATCAAACTGATTAATGACAAAAGATTAGAGAA 6553
Qy 5521 GAAAAATTAAGAAATTAAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAATTAAGAGTCAA 5580
Dh 6554 GAAAAATTAAGAAATTAAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAATTAAGAGTCAA 6613
Qy 5581 GTGGTCACAAAATGCCAGAGTCTTTCCGAGCAAGTGGACAGCTGCACTAGAGACTGGA 5640
Dh 6614 GTGGTCACAAAATGCCAGAGTCTTTCCGAGCAAGTGGACAGCTGCACTAGAGACTGGA 6673
Qy 5641 GTAGCAGTTAATTAATTAACAATACTCTGCAATATAAAAAATGTACTCAAAAT 5700
Dh 6674 GTAGCAGTTAATTAATTAACAATACTCTGCAATATAAAAAATGTACTCAAAAT 6733
Qy 5701 GTACGAAATGCTTGTGTAAATAAGCAAAATCTCATCTATTAAACAAATGGAATTTGA 5760
Dh 6734 GTACGAAATGCTTGTGTAAATAAGCAAAATCTCATCTATTAAACAAATGGAATTTGA 6793
Qy 5761 GCTGAGTTGGAGCTGGAGAGCTGGAGTGAAGTTCGTGTAGCACTGAAATTAAGATTGA 5820
Dh 6794 GCTGAGTTGGAGCTGGAGAGCTGGAGTGAAGTTCGTGTAGCACTGAAATTAAGATTGA 6853
Qy 5821 AATTAATGAGATGAGCAAAATTAATCAATGCAAAAATCACTGCGAAGGAAATGTCGAGTT 5880
Dh 6854 AATTAATGAGATGAGCAAAATTAATCAATGCAAAAATCACTGCGAAGGAAATGTCGAGTT 6913
Qy 5881 ATTACAGAGTCTGATCGGTAAATTTGCTAATTAATGACAGAAACAGTGTCTGAGTGGCCGT 5940
Dh 6914 ATTACAGAGTCTGATCGGTAAATTTGCTAATTAATGACAGAAACAGTGTCTGAGTGGCCGT 6973
Qy 5941 GCAGCAATAGAGCTCTCAACAGTGTGAATGAATTAACAGATCTCAAAAAGATTTGA 6000
Dh 6974 GCAGCAATAGAGCTCTCAACAGTGTGAATGAATTAACAGATCTCAAAAAGATTTGA 7033
Qy 6001 AAAAGATTCAAGTGTGCTAAGAGAAACGATGATTAATTAATCTCAAGGAGCA 6060
Dh 7034 AAAAGATTCAAGTGTGCTAAGAGAAACGATGATTAATTAATCTCAAGGAGCA 7093
Qy 6061 GTAGATTAAGTGTGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6120
Dh 7094 GTAGATTAAGTGTGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7153
Qy 6121 AAAAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6180
Dh 7154 AAAAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7213
Qy 6181 AAATCTTTTATGGCAAAATGCCGTGTGAGCAAGCCGAGATGGAGAACTGTTAAT 6240
Dh 7214 AAATCTTTTATGGCAAAATGCCGTGTGAGCAAGCCGAGATGGAGAACTGTTAAT 7273
Qy 6241 ATCAACAAAGTTATGAGAGAAACAGAACTCTTGTGAAAAATCTATATTAATTAATTAATTAAT 6300
Dh 7274 ATCAACAAAGTTATGAGAGAAACAGAACTCTTGTGAAAAATCTATATTAATTAATTAATTAAT 7333
Qy 6301 CATTATTTCTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6360
Dh 7334 CATTATTTCTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7393
Qy 6361 GTTGTGTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6420
Dh 7394 GTTGTGTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7453
Qy 6421 AAGCAAGAGTTGAGAAAAATCAACATGTCTGATGAAGTTTCGAGAGAAAGCTGAATTT 6480
Dh 7454 AAGCAAGAGTTGAGAAAAATCAACATGTCTGATGAAGTTTCGAGAGAAAGCTGAATTT 7513

Qy 6481 ACAGCAGATTTCTAAGCAAGAAATTTCTCTTTTGGAGTCCGAGTCCAGACAGCCGGGTA 6540
Dh 7514 ACAGCAGATTTCTAAGCAAGAAATTTCTCTTTTGGAGTCCGAGTCCAGACAGCCGGGTA 7573
Qy 6541 GAGCCGAGAGTGGAGAAACCGTTTCCGTAAATCAATTTGACGAAAGACGGAATTAAT 6600
Dh 7574 GAGCCGAGAGTGGAGAAACCGTTTCCGTAAATCAATTTGACGAAAGACGGAATTAATTAAT 7633
Qy 6601 GTGAGAGAGCAAGATTTTGGTAAATAAAGCTGAGTTTACAGCAAAAAGTTATAGTTCT 6660
Dh 7634 GTGAGAGAGCAAGATTTTGGTAAATAAAGCTGAGTTTACAGCAAAAAGTTATAGTTCT 7693
Qy 6661 GTTGCATTTGGAAATGCCAGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 6720
Dh 7694 GTTGCATTTGGAAATGCCAGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 7753
Qy 6721 GTGGCAGTTACCAAAAGATGAATCAAAACAGAGAGCAAGTGAATAATTTCTAATTTATG 6780
Dh 7754 GTGGCAGTTACCAAAAGATGAATCAAAACAGAGAGCAAGTGAATAATTTCTAATTTATG 7813
Qy 6781 ACTGCAAAACAGTTTGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6840
Dh 7814 ACTGCAAAACAGTTTGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7873
Qy 6841 TCAGCCGAGAGTGAATTTCTGACAGCCGAGTATCTGAGTGGTGTCTGCAATTAATTT 6900
Dh 7874 TCAGCCGAGAGTGAATTTCTGACAGCCGAGTATCTGAGTGGTGTCTGCAATTAATTT 7933
Qy 6901 GCAAAATTAAGTGAAGAAACAGATTCAGATCACTATTAATTAATTAATTAATTAATTAATTAAT 6960
Dh 7934 GCAAAATTAAGTGAAGAAACAGATTCAGATCACTATTAATTAATTAATTAATTAATTAATTAAT 7993
Qy 6961 GTAAAAAGCTCTTAATTAATTAATTTGCAATTTCTGACAGCCGAGTGGAGAGCCGAGTCTT 7020
Dh 7994 GTAAAAAGCTCTTAATTAATTAATTTGCAATTTCTGACAGCCGAGTGGAGAGCCGAGTCTT 8053
Qy 7021 GCAGCAGTTACCGAGTGGTGTCTGTAAACATTAATTAATTAATTAATTAATTAATTAATTAAT 7080
Dh 8054 GCAGCAGTTACCGAGTGGTGTCTGTAAACATTAATTAATTAATTAATTAATTAATTAATTAAT 8113
Qy 7081 CAAATTAATCTGATTTGAATTTGCAATTCGAGAGAAAGTAAATGTAACGCAAAAGAGAA 7140
Dh 8114 CAAATTAATCTGATTTGAATTTGCAATTCGAGAGAAAGTAAATGTAACGCAAAAGAGAA 8173
Qy 7141 AAAAATTAATTAAGCAAAACAGCAGAAATGAGAAATCGAGAGACGCAATTCGAGAGCAAT 7200
Dh 8174 AAAAATTAATTAAGCAAAACAGCAGAAATGAGAAATCGAGAGACGCAATTCGAGAGCAAT 8233
Qy 7201 GTCTTGTAAATTAATTTTGGAAACAGCTGAGAGAAATTAATTAATTAATTAATTAATTAATTAAT 7260
Dh 8234 GTCTTGTAAATTAATTTTGGAAACAGCTGAGAGAAATTAATTAATTAATTAATTAATTAATTAAT 8293
Qy 7261 ACAGAAAGTTTAAAACTTTAGACGAATTAACAAAGAAACAGATTAATTAATTAATTAATTAAT 7320
Dh 8294 ACAGAAAGTTTAAAACTTTAGACGAATTAACAAAGAAACAGATTAATTAATTAATTAATTAAT 8353
Qy 7321 GCTAGCAAAAATCTTAACATCAGAGTATTTCTACAGAAATCTTCTGTAAAGCG 7380
Dh 8354 GCTAGCAAAAATCTTAACATCAGAGTATTTCTACAGAAATCTTCTGTAAAGCG 8413
Qy 7381 GATAGAGAGATTAATCAGAGAGAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 7440
Dh 8414 GATAGAGAGATTAATCAGAGAGAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 8473
Qy 7441 GAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7500
Dh 8474 GAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8533
Qy 7501 GGAATTCAGAGTCTTCTCCGATCAGAAACAGTGGCAGTTTCAATTAATTAATTAATTAATTAAT 7560
Dh 8534 GGAATTCAGAGTCTTCTCCGATCAGAAACAGTGGCAGTTTCAATTAATTAATTAATTAATTAAT 8593

QY 7561 TC CGAGATTACTGTTGAAAAATCTTTTGTGAAAGCAGCTGAAAAAGTAATGTAGATCG 7620
Db 8594 TC CGAGATTACTGTTGAAAAATCTTTTGTGAAAGCAGCTGAAAAAGTAATGTAGATCG 8653
QY 7621 GATATTACAGGAAATGCTTCTTAA CAGATATCAAGCTCTGTAGAGCATTTGGGAATA 7680
Db 8654 GATATTACAGGAAATGCTTCTTAA CAGATATCAAGCTCTGTAGAGCATTTGGGAATA 8713
QY 7681 GGAGCTGCTATGACAGAAATTAATTTCTAATGGAAGATCAATATCATGATTTAAAAATCT 7740
Db 8714 GGAGCTGCTATGACAGAAATTAATTTCTAATGGAAGATCAATATCATGATTTAAAAATCT 8773
QY 7741 AAGCTATTAGAAAAAATATGTATGTTATTGTAAAAGATAAATCCGAATTTGAGACGGAA 7800
Db 8774 AAGCTATTAGAAAAAATATGTATGTTATTGTAAAAGATAAATCCGAATTTGAGACGGAA 8833
QY 7801 GCAAAAAGATTAA CCGTAGAGCCGTAGCTGCCGAGCCATTTATCTCAAAAGCAAAAGAT 7860
Db 8834 GCAAAAAGATTAA CCGTAGAGCCGTAGAGCCGTAGCTGCCGAGCCATTTATCTCAAAAGCAAAAGAT 8893
QY 7861 GAAATGAATTCAGAGGTTGAAAATTGAGAGATATTTCATGAGAAAATAGAGTAATCT 7920
Db 8894 GAAATGAATTCAGAGGTTGAAAATTGAGAGATATTTCATGAGAAAATAGAGTAATCT 8953
QY 7921 AGCCCTTCTAAAGAAATTTGAAAGAAATCAA TGTCAAAGTGAAAAAGAAAAACAGAGTG 7980
Db 8954 AGCCCTTCTAAAGAAATTTGAAAGAAATCAA TGTCAAAGTGAAAAAGAAAAACAGAGTG 9013
QY 7981 ACTGCTGAATCTCAAGAGACTCTGTAGAGACATGACAGGGGACAGAAATTTATCCGA 8040
Db 9014 ACTGCTGAATCTCAAGAGACTCTGTAGAGACATGACAGGGGACAGAAATTTATCCGA 9073
QY 8041 GCAAAAAGATGCGGAGAGCTCTTATTGAAAGTTAGTCAAAAATCCGGAAGAGTATTTT 8100
Db 9074 GCAAAAAGATGCGGAGAGCTCTTATTGAAAGTTAGTCAAAAATCCGGAAGAGTATTTT 9133
QY 8101 CATGCAGATTAATGTGAATATGGAAGCACATAAAAATGAAAGTAA CAGCAGTTCTAA 8160
Db 9134 CATGCAGATTAATGTGAATATGGAAGCACATAAAAATGAAAGTAA CAGCAGTTCTAA 9193
QY 8161 GCAGTAA CAGGTTCTGTATTTGGAGAGAGTTGGAGTACCAAGCAGAGCTACTGTGCA 8220
Db 9194 GCAGTAA CAGGTTCTGTATTTGGAGAGAGTTGGAGTACCAAGCAGAGCTACTGTGCA 9253
QY 8221 GGTAAACATATGATAGTAGAGTTGAGAGAAATTTGTTCAGACAAATGATTTGATGCA 8280
Db 9254 GGTAAACATATGATAGTAGAGTTGAGAGAAATTTGTTCAGACAAATGATTTGATGCA 9313
QY 8281 ATTTCTAAAGTAGAAGTTGATGAGAATAAGTAAGTCTTAATCTTCTGTAGTATCA 8340
Db 9314 ATTTCTAAAGTAGAAGTTGATGAGAATAAGTAAGTCTTAATCTTCTGTAGTATCA 9373
QY 8341 GGAATATGAGAGAGAAATTTGCCGAGCAGAGTGAATACTTCTACAGCACAAGTAATCT 8400
Db 9374 GGAATATGAGAGAGAAATTTGCCGAGCAGAGTGAATACTTCTACAGCACAAGTAATCT 9433
QY 8401 GAATCGTAGTTCGTTTACGAAAGCAAGATTATGAAATTAATGATTACCAAAAAATAT 8460
Db 9434 GAATCGTAGTTCGTTTACGAAAGCAAGATTATGAAATTAATGATTACCAAAAAATAT 9493
QY 8461 ATTTGAGAAGTCAATGCTCTTGCTTTAAATGATACAAAGATGAGCGAATATAGATCT 8520
Db 9494 ATTTGAGAAGTCAATGCTCTTGCTTTAAATGATACAAAGATGAGCGAATATATAGATCT 9553
QY 8521 TTAGCGGTAGCCGGTGTGATGCA CAAAGAA CAAACAAAGCATTTTAGAGTCAACAG 8580
Db 9554 TTAGCGGTAGCCGGTGTGATGCA CAAAGAA CAAACAAAGCATTTTAGAGTCAACAG 9613
QY 8581 TTAACTTCTACAACTGTAAATGAGAGAAAGTATCTCAACTTCGAGCAAAAGCTTTGGCT 8640
Db 9614 TTAACTTCTACAACTGTAAATGAGAGAAAGTATCTCAACTTCGAGCAAAAGCTTTGGCT 9673
QY 8641 AAAAAAGAAATTTATGAGAAATGTAAAAAGAACTGAGAGACCTTATGTCGAGACGGAAACA 8700

Db 9674 AAAAAAGAAATTTATGAGAAATGTAAAAAGAACTGAGAGACCTTATGTCGAGACGGAAACA 9733
QY 8701 GCAGCCGTTGAAAATTTATACAAAGATCTACAGAGCATTTGGTTGACAGAAATTTGGAA 8760
Db 9734 GCAGCCGTTGAAAATTTATACAAAGATCTACAGAGCATTTGGTTGACAGAAATTTGGAA 9793
QY 8761 ATTGAGATTAATTTAGAAACGATTTGCAGAGATATATA CAAATTTGAGAGTCAACGAGAC 8820
Db 9794 ATTGAGATTAATTTAGAAACGATTTGCAGAGATATATA CAAATTTGAGAGTCAACGAGAC 9853
QY 8821 GGAACCAAGAGAGCTCTTGTGGAAGAAATGATTTCTGTGAAAATTAACAATTTCAAGG 8880
Db 9854 GGAACCAAGAGAGCTCTTGTGGAAGAAATGATTTCTGTGAAAATTAACAATTTCAAGG 9913
QY 8881 GAAACAAAATCATCATTTGAGAGATTAAGCCAGAAATTTGTGAAACCGGAAGTAAATGTA 8940
Db 9914 GAAACAAAATCATCATTTGAGAGATTAAGCCAGAAATTTGTGAAACCGGAAGTAAATGTA 9973
QY 8941 GATGCTTTGAATGAATGATGTAGATCTACAGAGAAAGATGCTATGTGGAATTT 9000
Db 9974 GATGCTTTGAATGAATGATGTAGATCTACAGAGAAAGATGCTATGTGGAATTT 10033
QY 9001 GGTATTGGAATTTGTTGATGTAAATATGTGATTAAGAAAAATGTAGAACCAAAATCCGA 9060
Db 10034 GGTATTGGAATTTGTTGATGTAAATATGTGATTAAGAAAAATGTAGAACCAAAATCCGA 10093
QY 9061 AGACATGCTATTTGTAGAAACCTACAGAGAAACAGAAATTCAGACATTTCAAGAGCAAAA 9120
Db 10094 AGACATGCTATTTGTAGAAACCTACAGAGAAACAGAAATTCAGACATTTCAAGAGCAAAA 10153
QY 9121 GTAAATATTCTTTGGAAGAAAGAGACGCTCAGCTCAGCTGCAATTCGAATGTACACATT 9180
Db 10154 GTAAATATTCTTTGGAAGAAAGAGACGCTCAGCTCAGCTGCAATTCGAATGTACACATT 10213
QY 9181 TCCAATGAGATGATTTAAAAATTTGGCAAGCAGTATGCAATCTTCTCAATTAATAAC 9240
Db 10214 TCCAATGAGATGATTTAAAAATTTGGCAAGCAGTATGCAATCTTCTCAATTAATAAC 10273
QY 9241 AAAAAATTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9300
Db 10274 AAAAAATTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10333
QY 9301 GGGGTGCTGAAAGCAAGAGTGCAGAGACCAAGCGACAGTTAGTGAAGATCAATA 9360
Db 10334 GGGGTGCTGAAAGCAAGAGTGCAGAGACCAAGCGACAGTTAGTGAAGATCAATA 10393
QY 9361 AATAGAACTTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAACATCAATGTA 9420
Db 10394 AATAGAACTTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAACATCAATGTA 10453
QY 9421 TATGCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9480
Db 10454 TATGCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10513
QY 9481 GCCAAAGATCATGCTGACAGCTGCTTCCGCAACTGCACTATTTGAAAAATTAAGTAAAA 9540
Db 10514 GCCAAAGATCATGCTGACAGCTGCTTCCGCAACTGCACTATTTGAAAAATTAAGTAAAA 10573
QY 9541 TTTAATTAATGCGATCCGAAATTTAAAAATTAATCTGCAAGATTTGAAAGGAAAGCTAAT 9600
Db 10574 TTTAATTAATGCGATCCGAAATTTAAAAATTAATCTGCAAGATTTGAAAGGAAAGCTAAT 10633
QY 9601 AAAAAAAGCTGCTGATGATCTTAATAGTACAGTGTATACGAGATTAATTAATTAATTAAT 9660
Db 10634 AAAAAAAGCTGCTGATGATCTTAATAGTACAGTGTATACGAGATTAATTAATTAATTAAT 10693
QY 9661 TCTTCTGAAAAAGCATACAAAAATTTGACATCAATCAAAAGAGAGGAAAAAGGAAA 9720
Db 10694 TCTTCTGAAAAAGCATACAAAAATTTGACATCAATCAAAAGAGAGGAAAAAGGAAA 10753
QY 9721 AATATGA 9726

Db 10754 AATGA 10759

RESULT 3
AF312861

LOCUS
DEFINITION
Fusobacterium necrophorum Oppf (oppf) gene, partial cds; and lktB (lktB), lktA (lktA), and lktC (lktC) genes, complete cds.

ACCESSION
AF312861

VERSION
AF312861.3

KEYWORDS
GI:61676922

SOURCE
Fusobacterium necrophorum
Fusobacterium necrophorum
Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;

ORGANISM
Fusobacterium

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 12700)
Narayana, S.K., Nagara, T.G., Chengappa, M.M., and Stewart, G.C.
Cloning, sequencing, and expression of the leukotoxin gene from
Fusobacterium necrophorum

JOURNAL
PUBMED
11500416
Infect. Immun. 69 (9), 5447-5455 (2001)

REFERENCE
AUTHORS
TITLE
2 (bases 1 to 12700)
Oelke, A.M., Nagara, T.G., Wilkerson, M.J., and Stewart, G.C.
The leukotoxin operon of Fusobacterium necrophorum

JOURNAL
PUBMED
11500416
Unpublished

REFERENCE
AUTHORS
TITLE
3 (bases 1 to 12700)
Narayana, S., Stewart, G.C., Nagara, T.G., and Chengappa, M.M.
Direct Submision

JOURNAL
PUBMED
11500416
Submitted (12-OCT-2000) Diagnostic Medicine / Pathobiology, Kansas
State University, 1800 Denison Ave, Manhattan, KS 66506, USA

REFERENCE
AUTHORS
TITLE
4 (bases 1 to 12700)
Stewart, G.C., Nagara, S., Oelke, A.M., Nagara, T.G., and
Chengappa, M.M.
Direct Submision

JOURNAL
PUBMED
11500416
Submitted (23-FEB-2004) Diagnostic Medicine / Pathobiology, Kansas
State University, 1800 Denison Ave, Manhattan, KS 66506, USA

REFERENCE
AUTHORS
TITLE
5 (bases 1 to 12700)
Stewart, G.C., Nagara, S., Oelke, A.M., Nagara, T.G., and
Chengappa, M.M.
Direct Submision

JOURNAL
PUBMED
11500416
Submitted (22-MAR-2005) Diagnostic Medicine / Pathobiology, Kansas
State University, 1800 Denison Ave, Manhattan, KS 66506, USA

REMARK
COMMENT
On Mar 22, 2005 this sequence version replaced gi:42759853.

FEATURES
source
1..12700
/organism="Fusobacterium necrophorum"
/mol_type="genomic DNA"
/strain="A25"
/db_xref="taxon:859"
1..211
/gene="oppf"
1..211
/note="ABC transporter; nucleotide binding/ATPase protein"
/codon_start=2
/product="Oppf"
/protein_id="AA545303.1"
/db_xref="GI:42759853"
/translation="DLFEKIKKKRKLSCFISHDLFVNANICDRLVIMKEGKIEEGK
KEBIFRPAKKEVYQFLFLEVSMSFIF"
685..2337
/gene="lktB"
/note="lktB"
685..2337
/gene="lktA"
/note="putative transport protein"
/codon_start=1
/translation="LktB"
/product="LktB"
/protein_id="AAK27340.2"

/db_xref="GI:42759854"
/translation="WFHLKKELKTFVFFMILLSCPNILAEIGTQSTESDAGMLDRN
NRIPOGKLOKLIHSEKEREKOGIENFEKKKEBIDIKVDSVFLKKEIPSDILTER
EIVRIWEPIYIEKVTVDLITTVOKINELIYOKERYIVCRATLPKQKIONGIVNIIIE
KGTGDIILIQGNHSTREKYLIRERIPLEKORVSNFELBDSILTRFUNDSPLOIMASG
KPGTGTIVFQIYIEPKQOFPVFNIDNQKNTGLRWGLNINNSVGNRQSLTSL
LVTEGASIASSEYTPVPSKGTIKSLQHSVGLKHIQGLKHKITGNSYSYGVGIVHP1
LHBEKNKVSLEDMVKORTVTDILKLNKNRRLSKYTAGISISHBEDSVFTKQNI1T
KGRITPISGDARNTKYDMLIYOKNLKYNTLVTLKMGAGVSLSKLPSVBOIYVGA
YNRGYPENEGAEHGVFNALSKLVENKEFPVFLDAGSLHESAMQENRIPSSGP
GYKIRPLEKNKNIIVSMAPFWRKKLINSIVDSNRITYIYNHEP"
2353..12078
/gene="lktA"
2353..12078
/gene="lktC"
/note="Leukotoxin"
/codon_start=1
/translation="LktA"
/product="LktA"
/protein_id="AAK27341.1"
/db_xref="GI:13446804"
/translation="WSGIKNVQTRKRISSKVMMLGLLINTVYRANTITATE
NFGKIEKKDNDYDITNKKIGENAFNSFNPLTENNIVANLYFEKSKTGVNANLNF
VNGKIEVDGIIINGIRENKIGNLVFLSSEGAVGKQVINGSPHSI1PKODFFKAL
EAGKGVFNGLIIPVQGVKIPINPNSI1TVEGILNAVEGILYADIRLKDITAILKT
GITDPKLVNIDRINSGLTGDILKATKTKSGDIIILSHIDSPQKMGKSTVGRKIEB
VYKGTAKNISDVLNDGNIKISAKTATNRFIKKEGETVNTPLSLDVEASRV
NKGKIVTKNVIDTAEKVFYDALTITKLAKNSFSFVTSIPIINANGFLGILTSVSV
VIGKDAKEATEGKANHISYSGVATMTAAKSLPLKITLYEKAGKLISDAGTISA
KSNVITIEGVKSGRADITSKSNTIDASVSVTMDNSKNSALSVLTBENSISV
KILAKAGVSESTDVNVNSBAINSIRAAVKGLDPSGNGVAAANSINNASRILDVQ
YLHAKRLVNAHNITONSIVLQSGDLSQFPMNDHYVESGLKSLIDAIQRFQSGDS
VNEIKNLNLPVSVGATIANHNSNSVAIGSGRSLSSGVEGNNPALNKAOLRA
TTSSGVAVNRKEKKRLIGNAAVEFGNKNASVITIAHABLSVSGKIDNSENKIE
KNPSMAQSVIIDLKELKRAFGEKTEKTEYEPKDIJSEKLINASEKLDGKPELILN
GERMTIILPDDSGTALIELANYQMGKLEELPFGFKASLEGGLKILETINFT
GVGNANPHFTFSGANGERDVSIVGAAVSWVEQNYEKSVSGKAKLAALKADNLINKA
INKETVNLNIGLARSSTGSGANGELNORSKSAIIVEKKEALISEGNINDAL
NRLPVSAGIPNGSGGNAINMGASISGSISTRARVITIDBEATIKAKKILANIKNDTSV
NNPAAQSVISGRTVNRAPAKLALTQVTAIVTGGTISGAGLQNTGVANSVN
KIDNDVEASVDSISIEGANEINVIKADVSGSSDLAKEVQALNGDKKTLBDRGINTT
GNGYTTKEOLEKAKKKEGAVIVNALSVAGTDKSGAGVAINVTVKRFKALSGSNK
BAGEDKIHAKHVNVKASSTVYNNAAGSLAISKDAFSGSGSAMODLSIDTIAKDG
RISASLAVNNANSITGVNAGTIGASISLTVGGAANNITLHKTSALITGKPNFS
GNITRVNQLNDSHTIVSAGGAASIKOAGIGMVSVRGSDERBALVDSPEBGS
SFNVDAKQKTIINTIAGNANGKAGAVATVHNIIGKSVIAIVKNSKITTAANDDR
KNIINTAKDYMTNTIAGVGGAKGASVQASSTTLNKTNSHVDQDIDIDOLEEN
NGNKEKANNVLAENTSOVTVNATVLSGAGAAVAGAVAKIOTNDSAHIKNSON
VNALYKSHSISITIGIAGVAGAGVTSVNVKIVNNTIIELNIAKTIKAKNV
GYTSDAVIANYAGTVSGAALAGASVNEITGSPKAYVYKDSVIAKEFTDQIT
TQGVQKVDVRYAKNLTINEDLSQKRSKSNKGFPTNSATHTLUSILANNAAGSGAG
VAGTVNINKVGETEALVENSILNAKHSVKSQDITNSIGVGSVGVGAVGVASD
TNIIRKNTKTVGETTMSDEGFEAEITADSKOGISFSGVGAAGAAGVAVKIDES
NOFAGKTEVDVEAKILVKAIEITAKRSVVAIGNAVLAAGKAGIICAAVAVKIDES
NTRAVKSKIKMTNRKLDVIAENETKSGGIGSAGGILAAVGVSVNANIKVET
DIDHSTLHSTTVNPKALKINSISLTAGGAAGLAAVGVSVNNTNSVIARVHNS
DLTSVREKYNVTAKEEKNIKOTPAANGIGLAIIGAVLNNGTAVEDKNSBEGKTS
VAKTIDEVKEDQKKNVDAKTKILOSAGISTEDTSKADRGJTOGEGIAIVKTSOI1
GKAVDITTEDKNNITSTGIGTAGLASASGAVANNIKNSGTVLENSPVKAEKVN
RSDITGNVALTAYOGPVNIGAAVYAEIENSNGRNSISIKNSGLKNSIDVIAKQDE
LBAKAGLTVGAIVAAAGIISKAKNEMNSIVELEKSIPIENENVTSPSKIGEBINVK
EKENRYTASQASVGAAGACIISADVAGSSYLKVPSTKSGRSIPIADNVMALATKX
MKVTVASKAVTSVLGGVTVTAETAPAKTIVBEBGLPRTNRLNISKYBGLDED
KYTAKSVVSGGGIAGAGVNTASTQNSSTLSVNRKODYENDDTKTYISEVNLIA
LNDTKNEANIEBLAVAGVAGGTNKAFTSNSTLSTTVNGGAVNSQLRAALAKNEVY
NVKGCGALVGAETAVENTYKSTTGALVAGMEIGDKLETTIARDNTIIVRVNGDGTGK
GLVGNKIGISVKNITISGETKSI1EDKARIIVGTOSVNVVDALNEIDVDLOGSGVGIGI

```

gene
CDS
12075..12512
/gene="lkc"
12075..12512
/gene="lkc"
/codon_start=1
/transl_table=11
/product="Lkc"
/protein_id="AAS45304.1"
/db_xref="GI:4275985"
/translation="MKLRSEKSEFLKNSNITCFERBEVKDELTVYRSFMEVEGON
LPMVIMDNSITVINIROYAPVYIKDNTSEAVLSYINELNREYKFKFYVVEDADVCL
DSCVTSIAEEFPEMVYITLVNVLISHITHTYSTPMKKIIMSEK"

ORIGIN
Query Match      100.0%; Score 9726; DB 1; Length 12700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGAGGGGATCAAAAATAAGTTTCAGAGACAGAGAGAGAGATATCAATTTCTAAAAA  60
DB      2353  ATGAGCGGATCAAAAATAAGTTTCAGAGACAGAGAGAGAGATATCAATTTCTAAAAA  2412
QY      61  GTTTTATGATTTTGGGATTTGTTGATTAACACTATGACGGTGAAGGCTAATGATCAATC  120
DB      2413  GTTTTATGATTTTGGGATTTGTTGATTAACACTATGACGGTGAAGGCTAATGATCAATC  2472
QY      121  ACCGCGACTGAGAAATTTGGAACAAAATAGAAAAAAGAGATATGTTTATGACATTACT  180
DB      2473  ACCGCGACTGAGAAATTTGGAACAAAATAGAAAAAAGAGATATGTTTATGACATTACT  2532
QY      181  ACCAAACAGATTCAGAGGGAGAACGCTTTTAAACGTTTAAATAGATTTGCTTTAACAGAA  240
DB      2533  ACCAAACAGATTCAGAGGGAGAACGCTTTTAAACGTTTAAATAGATTTGCTTTAACAGAA  2592
QY      241  AATTAATATAGCAAAATCTATATTTTGGGAAAAAGAAATAGTACGGGGTAAATATCTTTT  300
DB      2593  AATTAATATAGCAAAATCTATATTTTGGGAAAAAGAAATAGTACGGGGTAAATATCTTTT  2652
QY      301  AACTTTGTCAATGAGAAAAATGGAAGTAGATGGGATTTATCAACGGAATTCGAGAAAAATAA  360
DB      2653  AACTTTGTCAATGAGAAAAATGGAAGTAGATGGGATTTATCAACGGAATTCGAGAAAAATAA  2712
QY      361  ATTGAGAGAAATTTATATTTCTTAAGCTCGGAAGGAGTGGCAGTAGAGAAAAATGAGATT  420
DB      2713  ATTGAGAGAAATTTATATTTCTTAAGCTCGGAAGGAGTGGCAGTAGAGAAAAATGAGATT  2772
QY      421  ATCAATGCTGATCTTTTCATTTCTATTTATCCAAAACAAGATGATTTTAAAGAGGCTTTG  480
DB      2773  ATCAATGCTGATCTTTTCATTTCTATTTATCCAAAACAAGATGATTTTAAAGAGGCTTTG  2832
QY      481  GAAAGAGCCAAACATGATAAGTTTAAATGGAATCATTCAGTAGATGAAAAAGTAAAA  540
DB      2833  GAAAGAGCCAAACATGATAAGTTTAAATGGAATCATTCAGTAGATGAAAAAGTAAAA  2892
QY      541  ATTCCATTGAATCCGAATGGAAGCATTAACGTTAGAAAGAAAAATCAATCTGTTGAAGC  600
DB      2893  ATTCCATTGAATCCGAATGGAAGCATTAACGTTAGAAAGAAAAATCAATCTGTTGAAGC  2952
QY      601  ATCGGTTTATGCGCGCGATATTAGATTGAAAGACTGCAATACTAANAACAGGAATT  660
DB      2953  ATCGGTTTATGCGCGCGATATTAGATTGAAAGACTGCAATACTAANAACAGGAATT  3012
QY      661  ACAGATTTTAAAAATTTAGTCAATATTAAGATCGAATAAATTCGTGTCGACCGAGAT  720
DB      3013  ACAGATTTTAAAAATTTAGTCAATATTAAGATCGAATAAATTCGTGTCGACCGAGAT  3072
QY      721  TTAAAAAGCTACCAAGCAAAAATCTGGAGATTAATTTCAAGCTCACTAAGATTCTCT  780
DB      721  TTAAAAAGCTACCAAGCAAAAATCTGGAGATTAATTTCAAGCTCACTAAGATTCTCTCT  780
```

```

DB      3073  TTAAAAAGCTACCAAGCAAAAATCTGGAGATTAATTTCTTCAAGCTCACATAGATTCTCT  3132
QY      781  CAAAAGCTATGGAAAAAAATTTCAACTGTTGGAAAAAGAAATAGAATATGTAANAAGA  840
DB      3133  CAAAAGCTATGGAAAAAAATTTCAACTGTTGGAAAAAGAAATAGAATATGTAANAAGA  3192
QY      841  AATACCAAGCAAAATATTGAATCTGATGCTGTATTGGAAAGCAGATGGAATATATAAATT  900
DB      3193  AATACCAAGCAAAATATTGAATCTGATGCTGTATTGGAAAGCAGATGGAATATATAAATT  3252
QY      901  AGTCGAAAGCTACAAAATGGAGAGATTTATAAGAAAGAGGGAAAAAGAACTTATATAC  960
DB      3253  AGTCGAAAGCTACAAAATGGAGAGATTTATAAGAAAGAGGGAAAAAGAACTTATATAC  3312
QY      961  ACTCCTTAAGTTTATCAGATGTAAGAGCTCCGTAGAAGTAAATTAAGAAAGAGTCATA  1020
DB      3313  ACTCCTTAAGTTTATCAGATGTAAGAGCTCCGTAGAAGTAAATTAAGAAAGAGTCATA  3372
QY      1021  GAAAAGAAATGTTGACATTACAGCTGAAAGCAAAAAGATTTCTATGATGCAACTTTAGTACT  1080
DB      3373  GAAAAGAAATGTTGACATTACAGCTGAAAGCAAAAAGATTTCTATGATGCAACTTTAGTACT  3432
QY      1081  AAGCTTGAAAGCACTCTTTAGCTTTGTTACAGGTTCTATTTCTCCATGCAATTTAAAT  1140
DB      3433  AAGCTTGAAAGCACTCTTTAGCTTTGTTACAGGTTCTATTTCTCCATGCAATTTAAAT  3492
QY      1141  GGAATTTTAAAGTTTATGACAAGATAGTCAAGTGTGTTATTTGAAAGAAATGCCAAAGTC  1200
DB      3493  GGAATTTTAAAGTTTATGACAAGATAGTCAAGTGTGTTATTTGAAAGAAATGCCAAAGTC  3552
QY      1201  GAAGCACAAGAAAGAAAGGCAATATTTCAATCTTAACAGTGAAGTAAAGCAACTATGGGA  1260
DB      3553  GAAGCACAAGAAAGAAAGGCAATATTTCAATCTTAACAGTGAAGTAAAGCAACTATGGGA  3612
QY      1261  GCAAGTACTTCTCCATTAATAAATTAACCAATTTATTTGGAGAAAGCCCAATGGAATACTT  1320
DB      3613  GCAAGTACTTCTCCATTAATAAATTAACCAATTTATTTGGAGAAAGCCCAATGGAATACTT  3672
QY      1321  CTCAGTATCGAGAGCGGATATATTTCTGCAAAAAGTAATTCCAATGTAACTATTGAAGA  1380
DB      3673  CTCAGTATCGAGAGCGGATATATTTCTGCAAAAAGTAATTCCAATGTAACTATTGAAGA  3732
QY      1381  GAAATTAATGCAAGGAGAAAGCAGATATTAATCTTCAAAATCTGAAATACTATTGATGCT  1440
DB      3733  GAAATTAATGCAAGGAGAAAGCAGATATTAATCTTCAAAATCTGAAATACTATTGATGCT  3792
QY      1441  TCTGTTCTGTTGGAACGATAGAGATTCCAATTAAGAGCTCTTTCAGTATTTGGTAGCG  1500
DB      3793  TCTGTTCTGTTGGAACGATAGAGATTCCAATTAAGAGCTCTTTCAGTATTTGGTAGCG  3852
QY      1501  GAAAGAGAAATTAATCTTCGCTCAAGATTGCTAAAGAGAGCAAAAAGTAAATCAAGAAACG  1560
DB      3853  GAAAGAGAAATTAATCTTCGCTCAAGATTGCTAAAGAGAGCAAAAAGTAAATCAAGAAACG  3912
QY      1561  GATGATGTAATGTAAGAGTAAGAGCGATTAATTCATTCGAGCTGCTGTAAGAGTGA  1620
DB      3913  GATGATGTAATGTAAGAGTAAGAGCGATTAATTCATTCGAGCTGCTGTAAGAGTGA  3972
QY      1621  TTGGGGGATAGTGGTAATGGGGTTGGGCTGCCAAATATTTCTAATCTAATAAGCTTCTCC  1680
DB      3973  TTGGGGGATAGTGGTAATGGGGTTGGGCTGCCAAATATTTCTAATCTAATAAGCTTCTCC  4032
QY      1681  CGTATGATGTAGATGGAATCTACATCCCAAGAAAGCACTAAATGTGGAAGCTCATAC  1740
DB      4033  CGTATGATGTAGATGGAATCTACATCCCAAGAAAGCACTAAATGTGGAAGCTCATAC  4092
QY      1741  ATTACTAAAAATAGTGTCTGCAAAACAGAGATCTGATTTGGAACTTCCAAAGTTATGAAT  1800
DB      4093  ATTACTAAAAATAGTGTCTGCAAAACAGAGATCTGATTTGGAACTTCCAAAGTTATGAAT  4152
QY      1801  GATCAGCTTTATGATCAGGTCATCTAAATCAATTTTAAATGCAATTAACAGCGGTTT  1860
DB      4153  GATCAGCTTTATGATCAGGTCATCTAAATCAATTTTAAATGCAATTAACAGCGGTTT  4212
```

QY 1861 GGAGGAGACGTGTCAATGAGGAATAAAGATAAGCTTAACGACTTATTTAGTCCGT 1920
DB 4213 GGAGGAGACAGTGTCAATGAGGAATAAAGATAAGCTTAACGACTTATTTAGTCCGT 4272
QY 1921 GTGTGCAACCATAGCAAAATCATATATATTTCTGCTCTGTGGCAATAGAGAGAGTGA 1980
DB 4273 GTGTGCAACCATAGCAAAATCATATATATTTCTGCTCTGTGGCAATAGAGAGAGTGA 4332
QY 1981 AGACTTTCTTCAGAGTGGAGAGTAATGTAGGGCAATTAAATGAAGCTCAAAATCTT 2040
DB 4333 AGACTTTCTTCAGAGTGGAGAGTAATGTAGGGCAATTAAATGAAGCTCAAAATCTT 4392
QY 2041 CGAGGACTACGTCAAGTGGAGAGTGGCTGTACGAAAGGAGAAAGAAAGAACTTAT 2100
DB 4393 CGAGGACTACGTCAAGTGGAGAGTGGCTGTACGAAAGGAGAAAGAAAGAACTTAT 4452
QY 2101 GGAATGCGAGCTTTTATGGAACCTAATAAAATGCTTCTGTGCAATGGCCAT 2160
DB 4453 GGAATGCGAGCTTTTATGGAACCTAATAAAATGCTTCTGTGCAATGGCCAT 4512
QY 2161 CATGCTGAATGCTATCGAGAGGAGAAATTTGATATCAACAGTGAATAATTTGATAT 2220
DB 4513 CATGCTGAATGCTATCGAGAGGAGAAATTTGATATCAACAGTGAATAATTTGATAT 4572
QY 2221 AAAAATCCTTCAAAAATGCGAAAGCTGTATTTAGTAATAATTTGAAAGAGCT 2280
DB 4573 AAAAATCCTTCAAAAATGCGAAAGCTGTATTTAGTAATAATTTGAAAGAGAGCT 4632
QY 2281 TTTGGAG 2340
DB 4633 TTTGGAG 4692
QY 2341 TTATGGAATGCAATTTTCAAGAAAAATTTGATGAGAAAAACCGAGCTTTTACTAATGGTGA 2400
DB 4693 TTATGGAATGCAATTTTCAAGAAAAATTTGATGAGAAAAACCGAGCTTTTACTAATGGTGA 4752
QY 2401 AGAATGACAAATTTCTTCGGATGGAATTTCAAAAACGAAATCTGTATAGAAATTTGA 2460
DB 4753 AGAATGACAAATTTCTTCGGATGGAATTTCAAAAACGAAATCTGTATAGAAATTTGA 4812
QY 2461 AACTATGCTCAGGAG 2520
DB 4813 AACTATGCTCAGGAG 4872
QY 2521 TTTTCAGAGAGATTTGAGTGAATTTGAAAGAACTTTGAAATTTTACAGAGAGAGAAAT 2580
DB 4873 TTTTCAGAGAGATTTGAGTGAATTTGAAAGAACTTTGAAATTTTACAGAGAGAGAAAT 4932
QY 2581 TATGCAAAATTTTCAACTTTTACCTTCTCCGAGCTAATGAGAGAGAGAGAGAGAGAG 2640
DB 4933 TATGCAAAATTTTCAACTTTTACCTTCTCCGAGCTAATGAGAGAGAGAGAGAGAGAG 4992
QY 2641 GTGGAGAGAGCTTTCTGTGGGTGAGAACGAGAGAAATTTAGCAAGGATTCCTGTGAGAAA 2700
DB 4993 GTGGAGAGAGCTTTCTGTGGGTGAGAACGAGAGAAATTTAGCAAGGATTCCTGTGAGAAA 5052
QY 2701 GGAGCTAAACTTCTGCAAAAAGAGATTTAAATATTTAAAGCTATCAATTAAGCAGAAAC 2760
DB 5053 GGAGCTAAACTTCTGCAAAAAGAGATTTAAATATTTAAAGCTATCAATTAAGCAGAAAC 5112
QY 2761 GTGAAATTTAGTGAATAATTTGACTTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
DB 5113 GTGAAATTTAGTGAATAATTTGACTTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5172
QY 2821 GGAAGATTAAATTTGAAAG 2880
DB 5173 GGAAGATTAAATTTGAAAG 5232
QY 2881 GAATTTATCAG 2940
DB 5233 GAATTTATCAG 5292

QY 2941 TCTTTAATGTGCTCAGGTGGAGATGCAATCAATGGAATGGAGTTATAGTGAGAGT 3000
DB 5293 TCTTTAATGTGCTCAGGTGGAGATGCAATCAATGGAATGGAGTTATAGTGAGAGT 5352
QY 3001 ATCAGTAAGGCAAGAGTTTCCATTGATGACGAAGCATATTTGAAAGCTAATTAAGAAAT 3060
DB 5353 ATCAGTAAGGCAAGAGTTTCCATTGATGACGAAGCATATTTGAAAGCTAATTAAGAAAT 5412
QY 3061 GCTTTAAACATTAAGATTAATCTTCTGTTTGAATGCTGCGGTTGACCGGGAATGGGA 3120
DB 5413 GCTTTAAACATTAAGATTAATCTTCTGTTTGAATGCTGCGGTTGACCGGGAATGGGA 5472
QY 3121 ACGAAAAATGCGGCGGTGCGGAGTTGCTGTGCGGTAATGATTAATTTCAAAACAA 3180
DB 5473 ACGAAAAATGCGGCGGTGCGGAGTTGCTGTGCGGTAATGATTAATTTCAAAACAA 5532
QY 3181 GCTTCATTTGAAGATTAATGACGAAGCAAAATTAATGATTAAGATTAAGATGATGA 3240
DB 5533 GCTTCATTTGAAGATTAATGACGAAGCAAAATTAATGATTAAGATTAAGATGATGA 5592
QY 3241 GTAACAGTAATCTGCGGAATCTTTAGAGTGAATGCAAAAACGACCGGAACTCAACAGT 3300
DB 5593 GTAACAGTAATCTGCGGAATCTTTAGAGTGAATGCAAAAACGACCGGAACTCAACAGT 5652
QY 3301 ATTTCTGTTGCGGAGAGATTAATTAAGGTTGAAGTAACCGAGTGAAGAAAAACCGAA 3360
DB 5653 ATTTCTGTTGCGGAGAGATTAATTAAGGTTGAAGTAACCGAGTGAAGAAAAACCGAA 5712
QY 3361 TCAG 3420
DB 5713 TCAG 5772
QY 3421 AATTAATTTAGAGATTAATGATTAATTTCAAGAAAAATTTGAAATTTTCACTTTCTGA 3480
DB 5773 AATTAATTTAGAGATTAATGATTAATTTCAAGAAAAATTTGAAATTTTCACTTTCTGA 5832
QY 3481 GGAGTAAAAAGAGCGGAGAGATTTCTCTCGAAAGTTTCTCACTCCGATTAAGAGAGAG 3540
DB 5833 GGAGTAAAAAGAGCGGAGAGATTTCTCTCGAAAGTTTCTCACTCCGATTAAGAGAGAG 5892
QY 3541 TCTTTCAAGTTGGAGCTTCTGGAAGTGTCTTTCAATTAATTTAAAAAGAGAGAGAG 3600
DB 5893 TCTTTCAAGTTGGAGCTTCTGGAAGTGTCTTTCAATTAATTTAAAAAGAGAGAGAG 5952
QY 3601 GCTGTGTAATGAGTAAGTAATTTGAAAGGAGCAATTAAGAGAGAGAGAGAGAGAG 3660
DB 5953 GCTGTGTAATGAGTAAGTAATTTGAAAGGAGCAATTAAGAGAGAGAGAGAGAGAG 6012
QY 3661 TCTTCTGATTTCACTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
DB 6013 TCTTCTGATTTCACTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6072
QY 3721 ATTGGAAGTGAAGATTAAGCAATCAATGCTGTTTGTGAGAGAGAGAGAGAGAGAG 3780
DB 6073 ATTGGAAGTGAAGATTAAGCAATCAATGCTGTTTGTGAGAGAGAGAGAGAGAGAG 6132
QY 3781 ATTCAAAAGTGAAGATTAAGCAATCAATGCTGTTTGTGAGAGAGAGAGAGAGAGAG 3840
DB 6133 ATTCAAAAGTGAAGATTAAGCAATCAATGCTGTTTGTGAGAGAGAGAGAGAGAGAG 6192
QY 3841 AAAAGTAATGCTTTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900
DB 6193 AAAAGTAATGCTTTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6252
QY 3901 AAAAGTAATGCTTTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3960
DB 6253 AAAAGTAATGCTTTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6312
QY 3961 GTGAACATTAAGAGTTTCTGCAAAATCAAGAAATTAATCAATGAGAGAGAGAGAG 4020
DB 6313 GTGAACATTAAGAGTTTCTGCAAAATCAAGAAATTAATCAATGAGAGAGAGAGAGAG 6372
QY 4021 CAAAAAATGAGAGTTGATGATCACTGCTTATCAAGCGAGACCAAGATGACAGAGAGCT 4080

|||||
Db 6373 CAAAAATGAGATGATGATCACTGCTTATCAAGCCGACACCCAGATGACAGAGCTTTA 6432
OY 4081 AATTTACAGCTGGAAAGTCAAAATGGAACGTGATGGGGCTACTGTGACGTGGCCAAATTA 4140
Db 6433 AATTTCAAGCTGGAAAGTCAAAATGGAACGTGATGGGGCTACTGTGACGTGGCCAAATTA 6492
OY 4141 AACCAACAAAGTAATGCTTCTATTAAGTGGGAGATATATCTAATCGTTAATCGAGCGAGC 4200
Db 6493 AACCAACAAAGTAATGCTTCTATTAAGTGGGAGATATATCTAATCGTTAATCGAGCGAGC 6552
OY 4201 GCAAAAAGCTCTTTTAGCAACCACTCAAGTGACTGTGACGTGACGACGGAGGAGACAAT 4260
Db 6553 GCAAAAAGCTCTTTTAGCAACCACTCAAGTGACTGTGACGTGACGACGGAGGAGACAAT 6612
OY 4261 AGTTCTGAGCGGAGTTAGGAAATATCAAGGGGCTGTTTCTGTCAATTAAGATTGACAT 4320
Db 6613 AGTTCTGAGCGGAGTTAGGAAATATCAAGGGGCTGTTTCTGTCAATTAAGATTGACAT 6672
OY 4321 GACGTGGAAGCTAGCGTTGATTAATCTTCCATCGAAGAGCTAATGAATCAATGTCATT 4380
Db 6673 GACGTGGAAGCTAGCGTTGATTAATCTTCCATCGAAGAGCTAATGAATCAATGTCATT 6732
OY 4381 GCCAAAGATGTCAAAGGAAGTTCTGATCTAGCAAAAAGAAATATCAGGCTTTACTAAATGGA 4440
Db 6733 GCCAAAGATGTCAAAGGAAGTTCTGATCTAGCAAAAAGAAATATCAGGCTTTACTAAATGGA 6792
OY 4441 AAAGATTAATAAATATTTAGAAAGATCGTGGATTAATATAGAGCTGGAATGGTATTATAGC 4500
Db 6793 AAAGATTAATAAATATTTAGAAAGATCGTGGATTAATATAGAGCTGGAATGGTATTATAGC 6852
OY 4501 AAGGAACAACTAGAAAAAGCAAAAGAAAGAAAGAGCGGCTCATTTGTAATCTGCTTTA 4560
Db 6853 AAGGAACAACTAGAAAAAGCAAAAGAAAGAAAGAGCGGCTCATTTGTAATCTGCTTTA 6912
OY 4561 TCGGTGCTGGAACCGATTAATTCGCTGGAGAGTACTATTTGCAATCAATCTGTTAA 4620
Db 6913 TCGGTGCTGGAACCGATTAATTCGCTGGAGAGTACTATTTGCAATCAATCTGTTAA 6972
OY 4621 AATTAATTTAAAGCAGAAATTTGAGTGAAGCAATTAAGGAACCGGAGAGGATTAATTCAT 4680
Db 6973 AATTAATTTAAAGCAGAAATTTGAGTGAAGCAATTAAGGAACCGGAGAGGATTAATTCAT 7032
OY 4681 GCGAAACATGTAATGTTGAGAGCAAAATCATCTACTGTTGTTGTGAATCGGCTTCTGA 4740
Db 7033 GCGAAACATGTAATGTTGAGAGCAAAATCATCTACTGTTGTTGTGAATCGGCTTCTGA 7092
OY 4741 CTTCCTATCAGCAAAATGCTTTTTCAGGAATGGGATCTGAGCATGCGAAGAATTATCA 4800
Db 7093 CTTCCTATCAGCAAAATGCTTTTTCAGGAATGGGATCTGAGCATGCGAAGAATTATCA 7152
OY 4801 AATGACAGATTTGCAAAAGTGAATAAGGAATTTCTGATGATTCTTTAATGTGAAC 4860
Db 7153 AATGACAGATTTGCAAAAGTGAATAAGGAATTTCTGATGATTCTTTAATGTGAAC 7212
OY 4861 GCAAAATTAATTCATTTCTGGGATGAATGTTGCGGGAACCATTCGCGTCTCTTTCTAGC 4920
Db 7213 GCAAAATTAATTCATTTCTGGGATGAATGTTGCGGGAACCATTCGCGTCTCTTTCTAGC 7272
OY 4921 GCGGTGAGAGCTGCTTTTGGGAATTAATCTTCAATATAAACCTCTGCTTTGATTACA 4980
Db 7273 GCGGTGAGAGCTGCTTTTGGGAATTAATCTTCAATATAAACCTCTGCTTTGATTACA 7332
OY 4981 GGAAGAGAGTAAATCTTTTATGTAAGAAATACAAAGTCAATGTACAGCTTTGAAT 5040
Db 7333 GGAAGAGAGTAAATCTTTTATGTAAGAAATACAAAGTCAATGTACAGCTTTGAAT 7392
OY 5041 GATTCATATTAACAAACGTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGGCTGGAATC 5100
Db 7393 GATTCATATTAACAAACGTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGGCTGGAATC 7452
OY 5101 GGAAGAAATGATCTGTCAATCGTGTCTGATGAACGGAAGCTTTAGTATGATTC 5160

|||||
Db 7453 GGAAGAAATGATATCTGTCAATCGTGGTTCTGATGAAACGGAAGCTTTAGTATGATTC 7512
OY 5161 GAGTTTGAAGAGTAAAGTTCCTTTTCAATGTAATGCAAAAGATCAAAATTAATTAACA 5220
Db 7513 GAGTTTGAAGAGTAAAGTTCCTTTTCAATGTAATGCAAAAGATCAAAATTAATTAACA 7572
OY 5221 ATTGCCGGAATTCGAAATGAGAGAAAGCGGCTGGAGTTTGAGCAACAGTTGCTATACA 5280
Db 7573 ATTGCCGGAATTCGAAATGAGAGAAAGCGGCTGGAGTTTGAGCAACAGTTGCTATACA 7632
OY 5281 AATATTTGAAAAACAATCGATTATAGCTTTATGTAATTTGTAATTTAACAACGCGCAAT 5340
Db 7633 AATATTTGAAAAACAATCGATTATAGCTTTATGTAATTTGTAATTTAACAACGCGCAAT 7692
OY 5341 GATCAAGATAGAAAAAATATCAATGTGACTGCAAAAAGATTATCTATAGCAACATCTATA 5400
Db 7693 GATCAAGATAGAAAAAATATCAATGTGACTGCAAAAAGATTATCTATAGCAACATCTATA 7752
OY 5401 GCAGTCGAGAGTTGGAGAGCAAAAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTAC 5460
Db 7753 GCAGTCGAGAGTTGGAGAGCAAAAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTAC 7812
OY 5461 TTGAATTAAGACAGTTTCTTCTCAATGTGATCAAACTGATATTTGACAAAGATTTAGAGAA 5520
Db 7813 TTGAATTAAGACAGTTTCTTCTCAATGTGATCAAACTGATATTTGACAAAGATTTAGAGAA 7872
OY 5521 GAAATTAATGGAATAATGGAATAAGCAAAATGTTAATGTTCTAGCTGAATAATACAGAGTCA 5580
Db 7873 GAAATTAATGGAATAATGGAATAAGCAAAATGTTAATGTTCTAGCTGAATAATACAGAGTCA 7932
OY 5581 GTGGTCACAAATGCGACAGTGTCTTTCGGAGCAAGTGAACAAGCTGACAGTGAAGCTGGA 5640
Db 7933 GTGGTCACAAATGCGACAGTGTCTTTCGGAGCAAGTGAACAAGCTGACAGTGAAGCTGGA 7992
OY 5641 GTAGAGATTAAATTAATTAACCAAAATCTTCTGACATATTAATAAATAGTACTAAAT 5700
Db 7993 GTAGAGATTAAATTAATTAACCAAAATCTTCTGACATATTAATAAATAGTACTAAAT 8052
OY 5701 GTAGCAAAATGCTTTGTTGTTAAAGCAAAATCTCATCTATTAATAAACAATTGGAATTGGA 5760
Db 8053 GTAGCAAAATGCTTTGTTGTTAAAGCAAAATCTCATCTATTAATAAACAATTGGAATTGGA 8112
OY 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGTGAACAAGTCTGTAGCAGTGAATTAAGTTGTA 5820
Db 8113 GCTGAGTTGAGAGCTGAGAGAGCTGAGTGAACAAGTCTGTAGCAGTGAATTAAGTTGTA 8172
OY 5821 AATAATTAAGATAGCAAAATTAATTAATCAATGCAAAATCACTGCGAAGGAAATGTGGAATT 5880
Db 8173 AATAATTAAGATAGCAAAATTAATTAATCAATGCAAAATCACTGCGAAGGAAATGTGGAATT 8232
OY 5881 ATTACAGAGTCTGATGCGGTATTTGCTAATTAATGACAGGAACAGTGTCTGAGTGGCCGCT 5940
Db 8233 ATTACAGAGTCTGATGCGGTATTTGCTAATTAATGACAGGAACAGTGTCTGAGTGGCCGCT 8292
OY 5941 GCAGCAATAGAGAGCTCAACAGTGTGAATGAATAATTAACAGATCTCAAAAAGCATATGTA 6000
Db 8293 GCAGCAATAGAGAGCTCAACAGTGTGAATGAATAATTAACAGATCTCAAAAAGCATATGTA 8352
OY 6001 AAAGATTTACAGATGATTGCTTAAGAGAAACAGATGATTATATTAATTAATCAAGGCGAA 6060
Db 8353 AAAGATTTACAGATGATTGCTTAAGAGAAACAGATGATTATTAATTAATCAAGGCGAA 8412
OY 6061 GTAGATAAGTGTGATTAAGTATTTCAAAAATCTTAATTAATTAAGGAAGCTTAATACAA 6120
Db 8413 GTAGATAAGTGTGATTAAGTATTTCAAAAATCTTAATTAATTAAGGAAGCTTAATACAA 8472
OY 6121 AAAAGAAAAATTAAGTATTAATAAAGAAATTTGTTACAAATAGTTCAGTACTATACCTTTA 6180
Db 8473 AAAAGAAAAATTAAGTATTAATAAAGAAATTTGTTACAAATAGTTCAGTACTATACCTTTA 8532
OY 6181 AAATCTTTATTTGGAATAGCGGCTGTGTTCAAGACAAAGCCGAGTGGCAGAACTGTTAAT 6240
Db 8533 AAATCTTTATTTGGAATAGCGGCTGTGTTCAAGACAAAGCCGAGTGGCAGAACTGTTAAT 8592

QY 6241 ATCAACAGTTTATGAGAAACAGAGCTCTGTGAGAAATTTCTATATTAATGCAAA 6300
DB 8593 ATCAACAGTTTATGAGAAACAGAGCTCTGTGAGAAATTTCTATATTAATGCAAA 8652
QY 6301 CATTTATCTGTAAATCAGAGATTAACAAGATTCATCGAGTAGTAGTTCTGTGGT 6360
DB 8653 CATTTATCTGTAAATCAGAGATTAACAAGATTCATCGAGTAGTAGTTCTGTGGT 8712
QY 6361 GTTGTGAAATGTAGAGTAGAGCTTCTGTATACCAATATTAATAAAGAAATAC 6420
DB 8713 GTTGTGAAATGTAGAGTAGAGCTTCTGTATACCAATATTAATAAAGAAATAC 8772
QY 6421 AAGAACAAGTTGAAAAATCTCAATGTCTGATGAAAGTTTGGAGAAAGCTGAAAT 6480
DB 8773 AAGAACAAGTTGAAAAATCTCAATGTCTGATGAAAGTTTGGAGAAAGCTGAAAT 8832
QY 6481 ACAGCAGATTTCTAACAAGAAATTTCTCTTTTGGAGTCGAGTCGACAGCCGGGTA 6540
DB 8833 ACAGCAGATTTCTAACAAGAAATTTCTCTTTTGGAGTCGAGTCGACAGCCGGGTA 8892
QY 6541 GGAGCCGAGTGCGAGAAACCGTTTCCGTAAATCAATTTGACAGAAAGCGAAGTAGAT 6600
DB 8893 GGAGCCGAGTGCGAGAAACCGTTTCCGTAAATCAATTTGACAGAAAGCGAAGTAGAT 8952
QY 6601 GTGGAAGAAGCAAAAGTTTGGTAAAAAAGCTGAGATTAACAGCAAAAAGTTTAGTTCT 6660
DB 8953 GTGGAAGAAGCAAAAGTTTGGTAAAAAAGCTGAGATTAACAGCAAAAAGTTTAGTTCT 9012
QY 6661 GTTGCAATTTGAAATGCGCAGTCGAGTCGAGTCGAAAGAGCTGCAATTTGAGACGA 6720
DB 9013 GTTGCAATTTGAAATGCGCAGTCGAGTCGAGTCGAAAGAGCTGCAATTTGAGACGA 9072
QY 6721 GTGCGAGTTTACCAAGATGAATCAACAGAGCAAGAGTAGAAAAATTTCTAAATTAATG 6780
DB 9073 GTGCGAGTTTACCAAGATGAATCAACAGAGCAAGAGTAGAAAAATTTCTAAATTAATG 9132
QY 6781 ACTCGAACAAGTTAGATGTAATAGCAAAAAATGAGATTAATCAGGTACTGGAAATCCGT 6840
DB 9133 ACTCGAACAAGTTAGATGTAATAGCAAAAAATGAGATTAATCAGGTACTGGAAATCCGT 9192
QY 6841 TCAGCCGAGTCGAAATTTCTGACGCGGAGTATCTGAGTGGTTTCTGCAATTAATTT 6900
DB 9193 TCAGCCGAGTCGAAATTTCTGACGCGGAGTATCTGAGTGGTTTCTGCAATTAATTT 9252
QY 6901 GCAAAATAGTAGAACAAGATTCGATCATAGTACTTTAACAATTTCTACTGATGTAAT 6960
DB 9253 GCAAAATAGTAGAACAAGATTCGATCATAGTACTTTAACAATTTCTACTGATGTAAT 9312
QY 6961 GTAAAAAGCTCTTAATTAATAATTTGAAATTCCTTGACAGCCGGTGAGAGCCCGAGTCTT 7020
DB 9313 GTAAAAAGCTCTTAATTAATAATTTGAAATTCCTTGACAGCCGGTGAGAGCCCGAGTCTT 9372
QY 7021 GGAAGCAGTTTACCGGAGTGGTTCTGTAAACAATTAATAGTTCTGTAGTAGCTCGAGTT 7080
DB 9373 GGAAGCAGTTTACCGGAGTGGTTCTGTAAACAATTAATAGTTCTGTAGTAGCTCGAGTT 9432
QY 7081 CAACAATACCTGATTTGACTTCCGTACAGAAAAATTAATGTAACGGCAAAAGAGAA 7140
DB 9433 CAACAATACCTGATTTGACTTCCGTACAGAAAAATTAATGTAACGGCAAAAGAGAA 9492
QY 7141 AAAAATATTAAGCAACAGACGAATTCAGAAATGAGAGACAGCAATCGAGCCAAAT 7200
DB 9493 AAAAATATTAAGCAACAGACGAATTCAGAAATGAGAGACAGCAATCGAGCCAAAT 9552
QY 7201 GTCTTGTAATTAATTTTGGAAACAGCTGTAGAAAGTAAGAAAAATTTCTGAAGAAAGGA 7260
DB 9553 GTCTTGTAATTAATTTTGGAAACAGCTGTAGAAAGTAAGAAAAATTTCTGAAGAAAGGA 9612
QY 7261 ACAGAAGTTTAAAACTTTAGACGAATTAACAAAGAACAGATTAATAAAGTAATGAT 7320
DB 9613 ACAGAAGTTTAAAACTTTAGACGAATTAACAAAGAACAGATTAATAAAGTAATGAT 9672

QY 7321 GCTAGAAAAAAATCTTACAATCAGCAGATATTTCTACAGAGATATCTTGTAAAAAGCG 7380
DB 9673 GCTAGAAAAAAATCTTACAATCAGCAGATATTTCTACAGAGATATCTTGTAAAAAGCG 9732
QY 7381 GATAGAGAGATATCTCAGGAGAGAGAAATTAAGCCATTGTGAGACTTCTGATTTATTT 7440
DB 9733 GATAGAGAGATATCTCAGGAGAGAGAAATTAAGCCATTGTGAGACTTCTGATTTATTT 9792
QY 7441 GAAAAAATGTAGATTAATACACAGAGACAAAGATTAATATCATCTTACTGTGTGGTTG 7500
DB 9793 GAAAAAATGTAGATTAATACACAGAGACAAAGATTAATATCATCTTACTGTGTGGTTG 9852
QY 7501 GGAACCTGACAGTCTTGCTTCGATCAGGAACAGTGCGAGTTACAAATTTAAAAAGAAAT 7560
DB 9853 GGAACCTGACAGTCTTGCTTCGATCAGGAACAGTGCGAGTTACAAATTTAAAAAGAAAT 9912
QY 7561 TCCGGAGTTACTGTGAAAAATCTTTGTGAAAGACGTGAAAAAGTAATGTTAGATCG 7620
DB 9913 TCCGGAGTTACTGTGAAAAATCTTTGTGAAAGACGTGAAAAAGTAATGTTAGATCG 9972
QY 7621 GATATTTACAGAAATGTTGCTTTAACAGATTCAGAGTCCGTGAGAGCATTTGGAAATA 7680
DB 9973 GATATTTACAGAAATGTTGCTTTAACAGATTCAGAGTCCGTGAGAGCATTTGGAAATA 10032
QY 7681 GGAGCTGCTATGACAGAAATTAATTTCTAATGGAAGATCAATATCATATTAATAATTTCT 7740
DB 10033 GGAGCTGCTATGACAGAAATTAATTTCTAATGGAAGATCAATATCATATTAATAATTTCT 10092
QY 7741 AAGCTATTAAGAAAAAATATTTAGATTTATGTAAGATTAATTCGAAATTTGAGACGGAA 7800
DB 10093 AAGCTATTAAGAAAAAATATTTAGATTTATGTAAGATTAATTCGAAATTTGAGACGGAA 10152
QY 7801 GCAAAAAGATTAACCGTAGAGACCGGTAGTCCCGGACCATTAATCTCAAAACCAAAAGAT 7860
DB 10153 GCAAAAAGATTAACCGTAGAGACCGGTAGTCCCGGACCATTAATCTCAAAACCAAAAGAT 10212
QY 7861 GAAATGAATTCAGAGTTGAAATTTGAGAGAGATATTTCAATGAAAGAAATAGATTAAT 7920
DB 10213 GAAATGAATTCAGAGTTGAAATTTGAGAGAGATATTTCAATGAAAGAAATAGATTAAT 10272
QY 7921 AGCCCTTCTAAGAAATTTGAGAGAGAAATCAATGTCAAAGTGAAAAAGAAAAACAGAGTG 7980
DB 10273 AGCCCTTCTAAGAAATTTGAGAGAGAAATCAATGTCAAAGTGAAAAAGAAAAACAGAGTG 10332
QY 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGTAGAGGCGAGAAATTAATTTCCGAA 8040
DB 10333 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGTAGAGGCGAGAAATTAATTTCCGAA 10392
QY 8041 GCAAAAAGATGCGGAGAGCTCTTAATTTGAAAGTTAGTACAAAAATCCGGAAGAAATTTTT 8100
DB 10393 GCAAAAAGATGCGGAGAGCTCTTAATTTGAAAGTTAGTACAAAAATCCGGAAGAAATTTTT 10452
QY 8101 CATGCAATATGTGAAATGGAAGCAACATTAATAATGAAATGTAACAGCAGTTCTTAA 8160
DB 10453 CATGCAATATGTGAAATGGAAGCAACATTAATAATGAAATGTAACAGCAGTTCTTAA 10512
QY 8161 GCAATACAGGTTCTGTATTTGAGAGAGTTGAGATCAACAGGACAGAGCTACTGCTCA 8220
DB 10513 GCAATACAGGTTCTGTATTTGAGAGAGTTGAGATCAACAGGACAGAGCTACTGCTCA 10572
QY 8221 GGTAAAACTATGTGAAAGTTGAGAGAGAAATTTGTTAGAACAAATGATGTAATGCA 8280
DB 10573 GGTAAAACTATGTGAAAGTTGAGAGAGAAATTTGTTAGAACAAATGATGTAATGCA 10632
QY 8281 ATTTCTAAGTAGAAGTTTGGATGATGAATTAAGTAACTGCTAAATCTTCTGTAGATCA 8340
DB 10633 ATTTCTAAGTAGAAGTTTGGATGATGAATTAAGTAACTGCTAAATCTTCTGTAGATCA 10692
QY 8341 GGAATAGAGAGAGAAATTCGAGACAGAGAGATGATATCTTCTACAGACAAAGTAATACT 8400
DB 10693 GGAATAGAGAGAGAAATTCGAGACAGAGAGATGATATCTTCTACAGACAAAGTAATACT 10752
QY 8401 GAATCCGTAGTTGCTTTAGAGAAAGCAAGATTTAGAAAAATTAATGATTAACAAAAAATAT 8460

```
Db 10753 GAATCCGATGTCGTTTACGAAGCAGATTATGAAAAATGATTACACAAAAAATAT 10812
Qy 8461 ATTTCAGAGTCATAGCTCTTGCTTAAATGATACAAAGAAAGAACCAATTAATACCT 8520
Db 10813 ATTTCAGAGTCATAGCTCTTGCTTAAATGATACAAAGAAAGAACCAATTAATACCT 10872
Qy 8521 TTAGCGGTAGCCGCTGTCATGACACAGAAACAAACCAATTCAGATCAACAAG 8580
Db 10873 TTAGCGGTAGCCGCTGTCATGACACAGAAACAAACCAATTCAGATCAACAAG 10932
Qy 8581 TTAACTTCTACACCTGTAAATGAGAGAAACGTATCTCAACTTCGCAAAAGCTTTGGCT 8640
Db 10933 TTAACTTCTACACCTGTAAATGAGAGAAACGTATCTCAACTTCGCAAAAGCTTTGGCT 10992
Qy 8641 AAAAAATGAAAAATTAAGAAATGTAAGAACTGAGAGACCTTAGTCGAGCGGAAACA 8700
Db 10993 AAAAAATGAAAAATTAAGAAATGTAAGAACTGAGAGACCTTAGTCGAGCGGAAACA 11052
Qy 8701 GCAGCCGTTGAAAAATTTATCAAGAGTACTACAGAGCATTGTTGCAAGAAATTTGGGA 8760
Db 11053 GCAGCCGTTGAAAAATTTATCAAGAGTACTACAGAGCATTGTTGCAAGAAATTTGGGA 11112
Qy 8761 ATTGAGATTAATTAAGAAACGATTGCAAGAGATTAACGATTGTAAGATCAACGAGAC 8820
Db 11113 ATTGAGATTAATTAAGAAACGATTGCAAGAGATTAACGATTGTAAGATCAACGAGAC 11172
Qy 8821 GGAACCAAGAGAGCTCTGTCGGAAGAAATGTTATTTCTGCAAAAAATTCATTTCAAGG 8880
Db 11173 GGAACCAAGAGAGCTCTGTCGGAAGAAATGTTATTTCTGCAAAAAATTCATTTCAAGG 11232
Qy 8881 GAAACAAAATCATCATTTGAAAGATTAAGCCAGAAATTTGTTGGAACCGGAAGTTAATGTA 8940
Db 11233 GAAACAAAATCATCATTTGAAAGATTAAGCCAGAAATTTGTTGGAACCGGAAGTTAATGTA 11292
Qy 8941 GATGCTTTGAATGAACCTTGATGATCTACAAAGAAAAAGTGTGCTATGTTGTAATTT 9000
Db 11293 GATGCTTTGAATGAACCTTGATGATCTACAAAGAAAAAGTGTGCTATGTTGTAATTT 11352
Qy 9001 GGTATTGGAAGAAATGTTGATGATTAATGATTAAGAAAAATGTAAGAACCAATTCGGA 9060
Db 11353 GGTATTGGAAGAAATGTTGATGATTAATGATTAAGAAAAATGTAAGAACCAATTCGGA 11412
Qy 9061 AGACATGCTATGTGAAACCTACTGAGAAAAACAAGATATCAAGCAATTTACAAGACAAA 9120
Db 11413 AGACATGCTATGTGAAACCTACTGAGAAAAACAAGATATCAAGCAATTTACAAGACAAA 11472
Qy 9121 GTAAATATTTCTGAAAAAGAGACGCTGCAAGCTGCAATATCGAATGTACACATTT 9180
Db 11473 GTAAATATTTCTGAAAAAGAGACGCTGCAAGCTGCAATATCGAATGTACACATTT 11532
Qy 9181 TCCCAATGAGATGATTAATTAATTTGGCAAGCAGTATGCAATCTTCTCAATTAATTAAC 9240
Db 11533 TCCCAATGAGATGATTAATTAATTTGGCAAGCAGTATGCAATCTTCTCAATTAATTAAC 11592
Qy 9241 AAAAAATGAAAAATTAATTAATTTAGCATCAAGTAAATGCAATGTAATGTAATTTGAT 9300
Db 11593 AAAAAATGAAAAATTAATTAATTTAGCATCAAGTAAATGCAATGTAATGTAATTTGAT 11652
Qy 9301 GGGGTGCTGGAAGCAGAGGTGCAAGAGCAGCAAGTAAATGTAATGTAATTAATTAATTA 9360
Db 11653 GGGGTGCTGGAAGCAGAGGTGCAAGAGCAGCAAGTAAATGTAATGTAATTAATTAATTA 11712
Qy 9361 AATAAGACTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAACCTCATGTA 9420
Db 11713 AATAAGACTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAACCTCATGTA 11772
Qy 9421 TATGCGGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9480
Db 11773 TATGCGGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11832
Qy 9481 GCCAAAGTCATGCTGACGCTCTTCGCAACTGCACTATGAAAAAATGAATGAATAAA 9540
```

```
Db 11833 GCCAAAGTCATGCTGACGCTCTTCGCAACTGCACTATGAAAAAATGAATGAATAAA 11892
Qy 9541 TTATATATGCGATCCGGAATTTAAAAATATATGCGCAAGTTGGAAGGAAAGCTAAT 9600
Db 11893 TTATATATGCGATCCGGAATTTAAAAATATATGCGCAAGTTGGAAGGAAAGCTAAT 11952
Qy 9601 AAAAAACGTGCGATAGATCTAATCAAGGTAGACTGATTAACGATTAATTAATTAATTAAT 9660
Db 11953 AAAAAACGTGCGATAGATCTAATCAAGGTAGACTGATTAACGATTAATTAATTAATTAAT 12012
Qy 9661 TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATTAAGAGAGGAGAAAAAGGAAA 9720
Db 12013 TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATTAAGAGAGGAGAAAAAGGAAA 12072
Qy 9721 AAATGA 9726
Db 12073 AAATGA 12078
```

```
RESULT 4
AR442022
LOCUS AR442022 2780 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 10 from patent US 6669940.
ACCESSION AR442022
VERSION AR442022.1 GI:42669173
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 2780)
Negara J. A., T. G., Stewart, G. C., Narayanan, S. K., and Chengappa, M. M.
Recombinant fuso bacterium necrophorum leukotoxin vaccine and
preparation thereof
Patent: US 6669940-A 10-30-DEC-2003;
JOURNAL Kansas University Research Foundation; Manhattan, KS
FEATURES
location/Qualifiers
1..2780
/mol_type="genomic DNA"
ORIGIN
```

```
Query Match 28.6%; Score 2780; DB 6; Length 2780;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 919 GGGAGATTTATTAAGAAAGGAGGAAAAAAGAACTTATACATCTCTTTAAGTTATCA 978
Db 1 GGGAGATTTATTAAGAAAGGAGGAAAAAAGAACTTATTAACATCTCTTTAAGTTATCA 60
Qy 979 GATGTGGAAGCTCCGTAAGATTAATTAAGAAAGCATAGGAAGAATGTTGACATTT 1038
Db 61 GATGTGGAAGCTCCGTAAGATTAATTAAGAAAGCATAGGAAGAATGTTGACATTT 120
Qy 1039 ACAGCTGAAGCAAGAAATTTCTATGATGCAACTTTAGTTACTAGCTTGCAAGCAGCT 1098
Db 121 ACAGCTGAAGCAAGAAATTTCTATGATGCAACTTTAGTTACTAGCTTGCAAGCAGCT 180
Qy 1099 TTTAGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAATGATTTTAAAGTTATTTAG 1158
Db 181 TTTAGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAATGATTTTAAAGTTATTTAG 240
Qy 1159 ACAAGTAAAGTCAGAGTGTGTTATTGGAAGAAAGTGCAGAAAGTGCAGAAAGCAAGAAAG 1218
Db 241 ACAAGTAAAGTCAGAGTGTGTTATTGGAAGAAAGTGCAGAAAGTGCAGAAAGCAAGAAAG 300
Qy 1219 GCAATATTCATTTCTTACAGTGAAGTAAGCAACTATAGGAGCAGCTACTTCTCATTA 1278
Db 301 GCAATATTCATTTCTTACAGTGAAGTAAGCAACTATAGGAGCAGCTACTTCTCATTA 360
Qy 1279 AAAATTAACCAATTTATATTTGAGAGAACCAATGGAATCTTCTAGATGAGAGCGGA 1338
Db 361 AAAATTAACCAATTTATATTTGAGAGAACCAATGGAATCTTCTAGATGAGAGCGGA 420
```

[illegible]

Db	1501	CCGGATGGAACCTTCAPAAAAACAGAACCTGTAATAGAAATTCGAAACTRATGTTCAAGGAGAA	1560
Qy	2479	ATGAAAAAATTAGAGNAAAAATTACCGAAAGGATTTAAAGCTTTTCAGAAAGATTGAGT	2538
Db	1561	ATGAAAAAATTAGAGNAAAAATTACCGAAAGGATTTAAAGCTTTTTCAGAAAGATTGAGT	1620
Qy	2539	GGACTGATTAAGAAACTTTTGAAATTTTACAGAGTAGAAATTAATGCAATTTTTCACACT	2598
Db	1621	GGACTGATTAAGAAACTTTTGAAATTTTACAGAGTAGAAATTAATGCAATTTTTCACACT	1680
Qy	2599	TTTACCTCTCCGAGCTAATGGAAGAAAGATGTTCTCTGAGGAGAGAGCTGTTTCG	2658
Db	1681	TTTACCTCTCCGAGCTAATGGAAGAAAGATGTTCTCTGAGGAGAGAGCTGTTTCG	1740
Qy	2659	TGGGTAGAACAGAGAAATTATAGCAAGTATCCGTTGAAAAGAGAGCTAAACTTGCTGCA	2718
Db	1741	TGGGTAGAACAGAGAAATTATAGCAAGTATCCGTTGAAAAGAGAGCTAAACTTGCTGCA	1800
Qy	2719	AAAAAGATTTAAATATPAAAGCTATCAATPAAACAGAAACAGTGAATTTTAGTTGCAAT	2778
Db	1801	AAAAAGATTTAAATATPAAAGCTATCAATPAAACAGAAACAGTGAATTTTAGTTGCAAT	1860
Qy	2779	ATTGACCTTGAGAACACAGTACATCCGGAAGTCAGTCCGAGGGAATTTAAATGTTCAA	2838
Db	1861	ATTGACCTTGAGAACACAGTACATCCGGAAGTCAGTCCGAGGGAATTTAAATGTTCAA	1920
Qy	2839	AGATCGAAAAATTCAGCTATCGTGAAGACTPAAAGAAAAGCTGAATTTATCAGAGAAAT	2898
Db	1921	AGATCGAAAAATTCAGCTATCGTGAAGACTPAAAGAAAAGCTGAATTTATCAGAGAAAT	1980
Qy	2899	ATTAAATGACATGCAATTGAAACAGACTTTTTCATGATAGCGGAGATCTTTTAATGCTGCTCA	2958
Db	1981	ATTAAATGACATGCAATTGAAACAGACTTTTTCATGATAGCGGAGATCTTTTAATGCTGCTCA	2040
Qy	2959	GCTGGGAATGCAATCAATGGAATGGGAAGTTATGTCGAGGTATCAGTAAGGCAAGGTT	3018
Db	2041	GCTGGGAATGCAATCAATGGAATGGGAAGTTATGTCGAGGTATCAGTAAGGCAAGGTT	2100
Qy	3019	TCCATTGTCAGAGCAATATTTGAAAGCTPATAAAAAAATTCGTTTAAACATPAGAAAT	3078
Db	2101	TCCATTGTCAGAGCAATATTTGAAAGCTPATAAAAAAATTCGTTTAAACATPAGAAAT	2160
Qy	3079	GATACTTGTGTTGGAATGCTGCCGTTCAAGCGGGAATCGGAACGAAAAATGCGCGCGCTC	3138
Db	2161	GATACTTGTGTTGGAATGCTGCCGTTCAAGCGGGAATCGGAACGAAAAATGCGCGCGCTC	2220
Qy	3139	GGGGTTCGTGTCGGGTAAATGATTAATGATTTTCAACAAAGCTCCATTGAAGATPAT	3198
Db	2221	GGGGTTCGTGTCGGGTAAATGATTAATGATTTTCAACAAAGCTCCATTGAAGATPAT	2280
Qy	3199	GACGAGAGCAAAAGTAAATATGATPAGAAATPAAAGATGAGATGAACTAACTCGGAA	3258
Db	2281	GACGAGAGCAAAAGTAAATATGATPAGAAATPAAAGATGAGATGAACTAACTCGGAA	2340
Qy	3259	TCTTTAGAGTAGATGCAAAAAACGACCGGAAACAATCAACAGTATTTCTGTTGCCGAGGA	3318
Db	2341	TCTTTAGAGTAGATGCAAAAAACGACCGGAAACAATCAACAGTATTTCTGTTGCCGAGGA	2400
Qy	3319	ATTAAATAGCTTGGAAAGTAAACGAGTAAACGAAATCCGAAATCAGAAAGAAACGAG	3378
Db	2401	ATTAAATAGCTTGGAAAGTAAACGAGTAAACGAAATCCGAAATCAGAAAGAAACGAG	2460
Qy	3379	GGAATTTTGGCAAAATTCGAAACAAAGTGACTCTGTAAAAAATTAATTCGGAATGAT	3438
Db	2461	GGAATTTTGGCAAAATTCGAAACAAAGTGACTCTGTAAAAAATTAATTCGGAATGAT	2520
Qy	3439	ATGATTCATTAACAGAAAAATTAACAATTAACATTTCTGAAGAGTAAAAAAGCGGAG	3498
Db	2521	ATGATTCATTAACAGAAAAATTAACAATTAACATTTCTGAAGAGTAAAAAAGCGGAG	2580
Qy	3499	AATCTTCTTGGAAGCTTTCATACCTCCGATAAAGACGCTCTTCAGTTTGGGAGCT	3558

Db 2581 AATCTTCCTCGACGTTTCTCATCTCCCGATAAAGACCGCTCTTTCAGTTGGAGCT 2640
Qy 3559 TCTGGAAGTGTCTTCTTCAATATATTAAGAAACATCTGCTGTGTAGATGAGTA 3618
Db 2641 TCTGGAAGTGTCTTCTTCAATATATTAAGAAACATCTGCTGTGTAGATGAGTA 2700
Qy 3619 AAGATTAATTTGAAGGAGCAATAAAGGTAGAGTGACTCTTCTGATTTCTACTTT 3678
Db 2701 AAGATTAATTTGAAGGAGCAATAAAGGTAGAGTGACTCTTCTGATTTCTACTTT 2760
Qy 3679 GTTGGAGCATGGGGCGGATC 3698
Db 2761 GTTGGAGCATGGGGCGGATC 2780

RESULT 5
AR442025
LOCUS AR442025 2322 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 13 from patent US 6669940.
ACCESSION AR442025
VERSION AR442025.1 GI:42669176
KEYWORDS
SOURCE
ORGANISM
Unclassified.

REFERENCE
1 (bases 1 to 2322)
Nagaraj,A.,T.G., Stewart,G.C., Narayanan,S.K. and Chengappa,M.M.
AUTHORS Recombinant fimbriated bacterium necrophorum leukotoxin vaccine and
TITLE Preparation thereof
JOURNAL Patent: US 6669940-A 13 30-DEC-2003;
Kansas University Research Foundation; Manhattan, KS
FEATURES
source
1..2322
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 23.9%; Score 2322; DB 6; Length 2322;
Best Local Similarity 100.0%; Pred. No. 4e-296;
Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7405 GGAATTTAAAGCATTGTGAAGCTTCTGATATTTATTTGAAAAAATGTAGATTTTCAACA 7464
Db 1 GGAATTTAAAGCATTGTGAAGCTTCTGATATTTATTTGAAAAAATGTAGATTTTCAACA 60
Qy 7465 GAGGACAAGATTAATATCACTTCTAGTGTGGGAACTGCAAGTCTTCTCCGA 7524
Db 61 GAGGACAAGATTAATATCACTTCTAGTGTGGGAACTGCAAGTCTTCTCCGA 120
Qy 7525 TCAGGACAAGTGTGCAATTAATAAAGAAATTCGAGATTACTGTTGAAATTTCT 7584
Db 121 TCAGGACAAGTGTGCAATTAATAAAGAAATTCGAGATTACTGTTGAAATTTCT 180
Qy 7585 TTTGTGAAAAGCAGCTGAAAAAGTAAATGTTAGATCGATATTAACAGAAATGTTCTTTA 7644
Db 181 TTTGTGAAAAGCAGCTGAAAAAGTAAATGTTAGATCGATATTAACAGAAATGTTCTTTA 240
Qy 7645 ACAGCATATCAAGTCTCTAGAGCATTTGGAAATGAGAGTCCCTATGCAATTTAAT 7704
Db 241 ACAGCATATCAAGTCTCTAGAGCATTTGGAAATGAGAGTCCCTATGCAATTTAAT 300
Qy 7705 TCTAAATGGAAGTCAATATCAATATTAATAATTTCTAAGCTTTAGAAAAATATTTGAT 7764
Db 301 TCTAAATGGAAGTCAATATCAATATTAATAATTTCTAAGCTTTAGAAAAATATTTGAT 360
Qy 7765 GTTATTTGTAAGATTAATCGAATTTGAGAGCGAAACAAAGATTAACCGTAGAGCG 7824
Db 361 GTTATTTGTAAGATTAATCGAATTTGAGAGCGAAACAAAGATTAACCGTAGAGCG 420
Qy 7825 GTAGCTGCCGAGCCATTATCTCAAAAGCAAGAAATGAAATGAAATTCAGAGGTTGAAAT 7884
Db 421 GTAGCTGCCGAGCCATTATCTCAAAAGCAAGAAATGAAATGAAATTCAGAGGTTGAAAT 480

Qy 7885 GAGAGAGTATTTTCAATGAAGAAATAGAGTAATCAAGCCCTTCTTAAGAAATTTGAAAG 7944
Db 481 GAGAGAGTATTTTCAATGAAGAAATAGAGTAATCAAGCCCTTCTTAAGAAATTTGAAAG 540
Qy 7945 GAATCAATGTCCAAAGTGGAAAAAGAAACAGAGTACTGCTGAATCTCAAGAGCTTCT 8004
Db 541 GAATCAATGTCCAAAGTGGAAAAAGAAACAGAGTACTGCTGAATCTCAAGAGCTTCT 600
Qy 8005 GTAGAGCAGTACGAGGCGAGAAATTTATTCGAAAGCAAAAGATGCGGAAAGCTTTAT 8064
Db 601 GTAGAGCAGTACGAGGCGAGAAATTTATTCGAAAGCAAAAGATGCGGAAAGCTTTAT 660
Qy 8065 TTGAAAGTTAGTACAAATCCGGAAGAAATTTTTCATGAGATTAATGTAATTTGAA 8124
Db 661 TTGAAAGTTAGTACAAATCCGGAAGAAATTTTTCATGAGATTAATGTAATTTGAA 720
Qy 8125 GCAACACATTAATGAAGTAAACAGCAAGTTTCTAAGCAGTAAACAGGTTCTGATTTGGA 8184
Db 721 GCAACACATTAATGAAGTAAACAGCAAGTTTCTAAGCAGTAAACAGGTTCTGATTTGGA 780
Qy 8185 GGAAGTTGAGTCAACCAAGCAGAAAGTCTGCTGAGGTTAAACATAAGTAAAGTTGAG 8244
Db 781 GGAAGTTGAGTCAACCAAGCAGAAAGTCTGCTGAGGTTAAACATAAGTAAAGTTGAG 840
Qy 8245 GAAAGAAATTTGTTCAAGAAACAAATGATTAATGCAATTTCTTAAGTAAAGTTGAT 8304
Db 841 GAAAGAAATTTGTTCAAGAAACAAATGATTAATGCAATTTCTTAAGTAAAGTTGAT 900
Qy 8305 GAAATTAAGTAACTGCTTAATTTCTTCTGATTAATGAGAAATGAGAGAGAAATTCGCGA 8364
Db 901 GAAATTAAGTAACTGCTTAATTTCTTCTGATTAATGAGAAATGAGAGAGAAATTCGCGA 960
Qy 8365 GCAGAGTGAATTAATCTTCAACAGCAAAAGTAACTGTAATCCGATTCGTTTACGAAG 8424
Db 961 GCAGAGTGAATTAATCTTCAACAGCAAAAGTAACTGTAATCCGATTCGTTTACGAAG 1020
Qy 8425 CAAGATTAAGAAATTAATGATTAACAAAAAAATATATTTGAGAGTCAATGCTTGTCT 8484
Db 1021 CAAGATTAAGAAATTAATGATTAACAAAAAAATATATTTGAGAGTCAATGCTTGTCT 1080
Qy 8485 TTTAAATGATTAAGAAATGAAGCCAAATTAAGATTTTAAAGGTTAGCCGTTGTGATCA 8544
Db 1081 TTTAAATGATTAAGAAATGAAGCCAAATTAAGATTTTAAAGGTTAGCCGTTGTGATCA 1140
Qy 8545 CAAGGACAAACAAAGCAATTTACGAGTCAACAAAGTTAACTTCAACAGTAAATGGA 8604
Db 1141 CAAGGACAAACAAAGCAATTTACGAGTCAACAAAGTTAACTTCAACAGTAAATGGA 1200
Qy 8605 GAAAACGATCTCAACTTCTGCAAAAGCTTTGCTAATAAATGAAAATTAATGAAATGTA 8664
Db 1201 GAAAACGATCTCAACTTCTGCAAAAGCTTTGCTAATAAATGAAAATTAATGAAATGTA 1260
Qy 8665 AAAGAACTGAGAGAGCTTATGTCGAGCGGAAACAGACGCTTGAATAATTAACAAG 8724
Db 1261 AAAGAACTGAGAGAGCTTATGTCGAGCGGAAACAGACGCTTGAATAATTAACAAG 1320
Qy 8725 AGTACTACAGAGCAATGTTGTCAGGAAATTTGAGAAATTTGAGATTAATTTGAAACGAT 8784
Db 1321 AGTACTACAGAGCAATGTTGTCAGGAAATTTGAGAAATTTGAGATTAATTTGAAACGAT 1380
Qy 8785 GCAAGAGATTAATACGATTTGTAAGATCAACGAGACGGAACCAAGAGAGTCTTGTCCGA 8844
Db 1381 GCAAGAGATTAATACGATTTGTAAGATCAACGAGACGGAACCAAGAGAGTCTTGTCCGA 1440
Qy 8845 AAAGATGATTTTCTGTGAAATAATCAATTTTCAGGGGAAACAAATTCATCTGTAAGAT 8904
Db 1441 AAAGATGATTTTCTGTGAAATAATCAATTTTCAGGGGAAACAAATTCATCTGTAAGAT 1500
Qy 8905 AAAGCAGAAATTTGTTGGAACCGGAAGTGAATTAATGTAAGTCTTGAAGAACTTGAATGA 8964
Db 1501 AAAGCAGAAATTTGTTGGAACCGGAAGTGAATTAATGTAAGTCTTGAAGAACTTGAATGA 1560
Qy 8965 GATCTACAGGAAAAAGTGTGCTATGTGGAATTTGTAATTTGAAATGTTGAATGTAAT 9024

Qy	4573	ACGATTAATCCGCTGAGAGAGTACGTAATGCAATACCTGTTAAAAATTAATTTAA	4632
Db	1021	ACGGATTAATCCGCTGAGAGAGTACGTAATGCAATACCTGTTAAAAATTTAA	1080
Qy	4633	GCAGATTAAGTGAAGCAATTAAGGAGCGGAGAGATAAAATTCAGCCGAACATGTA	4692
Db	1081	GCAGATTAAGTGAAGCAATTAAGGAGCGGAGAGATAAAATTCAGCCGAACATGTA	1140
Qy	4693	AATGTGAGAGCAAAATCATCTACTGTTGTGTGAATGCGGCTTCGACCTTGCTATCAGC	4752
Db	1141	AATGTGAGAGCAAAATCATCTACTGTTGTGTGAATGCGGCTTCGACCTTGCTATCAGC	1200
Qy	4753	AAAGATGCTTTTTCAGAGAAATGAGATCTGAGCATGCGAAGACTTATCAAAATGACAGAT	4812
Db	1201	AAAGATGCTTTTTCAGAGAAATGAGATCTGAGCATGCGAAGACTTATCAAAATGACAGAT	1260
Qy	4813	GCAAAAGTGAATTAAGAGAAATTTCTGTGATTCCTTAAATGTGAACCAATTAATTC	4872
Db	1261	GCAAAAGTGAATTAAGAGAAATTTCTGTGATTCCTTAAATGTGAACCAATTAATTC	1320
Qy	4873	ATTCTTGGGGTGAATGTTGCGGGAACCATTTGCCGGTCTCTTCTACGGCGGTAGAGCT	4932
Db	1321	ATTCTTGGGGTGAATGTTGCGGGAACCATTTGCCGGTCTCTTCTACGGCGGTAGAGCT	1380
Qy	4933	GCTTTTGGCAATTAATCTCTTCAATAAATTAATTAATTAATTAATTAATTAATTAAT	4992
Db	1381	GCTTTTGGCAATTAATCTCTTCAATAAATTAATTAATTAATTAATTAATTAATTAAT	1440
Qy	4993	AATCCTTTAGTGAAGAAATCAAAAGTCATGTAACAAGCTTTGAATTCCTCATATT	5052
Db	1441	AATCCTTTAGTGAAGAAATCAAAAGTCATGTAACAAGCTTTGAATTCCTCATATT	1500
Qy	5053	ACAAAGCTTTCTGCTGAGAGCGCTGCAAGTATTAAGACAGGCTGGAATCGAGAAATGTA	5112
Db	1501	ACAAAGCTTTCTGCTGAGAGCGCTGCAAGTATTAAGACAGGCTGGAATCGAGAAATGTA	1560
Qy	5113	TCTGTCAATCGTGTCTGTGATGAACGAAAGCTTTAGTGTGATTCGAGTTGAAGA	5172
Db	1561	TCTGTCAATCGTGTCTGTGATGAACGAAAGCTTTAGTGTGATTCGAGTTGAAGA	1620
Qy	5173	GTAAGTCTCTTCAATGTAATGCAAAAGTCAAAATTAATTAATTAATTAATTAATTAAT	5232
Db	1621	GTAAGTCTCTTCAATGTAATGCAAAAGTCAAAATTAATTAATTAATTAATTAATTAAT	1680
Qy	5233	GCAAAATGAGAGAAAGCGGCTGAGTGAAGCAAGTTGCTCATACAAATTTGGAANA	5292
Db	1681	GCAAAATGAGAGAAAGCGGCTGAGTGAAGCAAGTTGCTCATACAAATTTGGAANA	1740
Qy	5293	CAATCAATTATGCTATTGTAAAAAAGTAATAATTAACACGCGCAATGATCAGATGA	5352
Db	1741	CAATCAATTATGCTATTGTAAAAAAGTAATAATTAACACGCGCAATGATCAGATGA	1800
Qy	5353	AAAAATTAATCAATGTAATGCAAAAGTATTAATTAATTAATTAATTAATTAATTAAT	5412
Db	1801	AAAAATTAATCAATGTAATGCAAAAGTATTAATTAATTAATTAATTAATTAATTAAT	1860
Qy	5413	GAGAGAGCAAAAGAGCCCTGCTGCAAGAGCTTCTGCAAGTACTACCTGAAATGAACA	5472
Db	1861	GAGAGAGCAAAAGAGCCCTGCTGCAAGAGCTTCTGCAAGTACTACCTGAAATGAACA	1920
Qy	5473	GTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGAGAAATTAATGA	5532
Db	1921	GTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGAGAAATTAATGA	1980
Qy	5533	AATTAAGAGAAAGCAAAATGTTAATGTTGAGTGAATAATGAGTCAAGTGTCAAAAT	5592
Db	1981	AATTAAGAGAAAGCAAAATGTTAATGTTGAGTGAATAATGAGTCAAGTGTCAAAAT	2040
Qy	5593	GCGACAGTGTCTTCCGAGAGCAAGTGAACAAGCTGAGTGAAGCTGAGTACAGTTAAT	5652
Db	2041	GCGACAGTGTCTTCCGAGAGCAAGTGAACAAGCTGAGTGAAGCTGAGTACAGTTAAT	2100
Qy	5653	AAAAATTACAAAAATTAATCTTGTGACATATATAAAAAATAGTAC	5693

Db	2101	AAAAATTACAAAAATTAATCTTGTGCACATATATAAAAAATAGTAC	2141
RESULT 7			
LOCUS	AR442024	1887 bp	DNA linear PAT 20-FEB-2004
DEFINITION	Sequence 12 from patent US 6669940.		
ACCESSION	AR442024		
VERSION	AR442024.1 GI:42669175		
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1887)		
AUTHORS	Negareja,T.G., Stewart,G.C., Narayanan,S.K. and Chengappa,M.M.		
TITLE	Recombinant fimbriobacterium necrophorum leukotoxin vaccine and preparation thereof		
JOURNAL	Patent: US 6669940-A 12 30-DEC-2003;		
FEATURES	Kansas University Research Foundation; Manhattan, KS		
source	1..1887		
ORIGIN	/organism="unknown"		
	/mol_type="genomic DNA"		
Query Match 19.4%; Score 1887; DB 6; Length 1887;			
Best Local Similarity 100.0%; Pred. No. 7.4e-239;			
Matches 1887; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	5624	CTGCAGTAGAGAGCTGAGTACAGCTTAATTAATTAATTAATTAATTAATTAATTAAT	5683
Db	1	CTGCAGTAGAGAGCTGAGTACAGCTTAATTAATTAATTAATTAATTAATTAATTAAT	60
Qy	5684	AAATAGTACTCAAAATGTAGAAATGCTTTGGTGTAAAAAGCAAAATCTCATCTATTA	5743
Db	61	AAATAGTACTCAAAATGTAGAAATGCTTTGGTGTAAAAAGCAAAATCTCATCTATTA	120
Qy	5744	AAACAATTGGAATGGAGCTGAGTGGAGTGGAGCTGAGAGAGCTGAGTGAACAGGTTCTGAG	5803
Db	121	AAACAATTGGAATGGAGCTGAGTGGAGTGGAGCTGAGAGAGCTGAGTGAACAGGTTCTGAG	180
Qy	5804	CAGTGAATTAAGTTGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	5863
Db	181	CAGTGAATTAAGTTGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	240
Qy	5864	AGGGAATGTGGAGTTTACAGAGTCTGATGCGGTAAATGCTAATTATGCAAGAACAG	5923
Db	241	AGGGAATGTGGAGTTTACAGAGTCTGATGCGGTAAATGCTAATTATGCAAGAACAG	300
Qy	5924	TGTCTGAGTGGCCCGTGCACCAATAGGAGGCTCAACAGGTGAAATGAATTAACAGAT	5983
Db	301	TGTCTGAGTGGCCCGTGCACCAATAGGAGGCTCAACAGGTGAAATGAATTAACAGAT	360
Qy	5984	CTACAAAAGCATATGTAATAAATTTCTACAGATTTGCTAAAGAGAAACAGATGATTATA	6043
Db	361	CTACAAAAGCATATGTAATAAATTTCTACAGATTTGCTAAAGAGAAACAGATGATTATA	420
Qy	6044	TTACTACTCAAGGCGAGTGAATTAAGTGTGATTAAGTATTCAAAAATCTTAATTA	6103
Db	421	TTACTACTCAAGGCGAGTGAATTAAGTGTGATTAAGTATTCAAAAATCTTAATTA	480
Qy	6104	ACGAAGACTTATCACAAAAAAGAAAAAATTAATTAATTAATTAATTAATTAATTAAT	6163
Db	481	ACGAAGACTTATCACAAAAAAGAAAAAATTAATTAATTAATTAATTAATTAATTAAT	540
Qy	6164	CAGCTACTCATCTTTAAATCTTTATTTGGAATAATGCGCTGTTTCAAGAACGCGAG	6223
Db	541	CAGCTACTCATCTTTAAATCTTTATTTGGAATAATGCGCTGTTTCAAGAACGCGAG	600
Qy	6224	TGGCAGGAACGTTTAATTAATTAACAAGTTTAATGAGAAACGAAGCTCTTGTAGAAATTT	6283
Db	601	TGGCAGGAACGTTTAATTAATTAACAAGTTTAATGAGAAACGAAGCTCTTGTAGAAATTT	660

QY 6284 CTATATTAATGCAAAACATTATCTGTAAATCAGAGATTACAGAAATTCATCGAG 6343
DB 661 CTATATTAATGCAAAACATTATCTGTAAATCAGAGATTACAGAAATTCATCGAG 720
QY 6344 TAGAGGTTCTGTGGTGTGTGTGAAATGAGAGTAGAGCTTCTTCTGATACAAAT 6403
DB 721 TAGAGGTTCTGTGGTGTGTGTGAAATGAGAGTAGAGCTTCTTCTGATACAAAT 780
QY 6404 TTATATTAATGCAAAACATTATCTGTAAATCAGAGATTACAGAAATTCATCGAG 6463
DB 781 TTATATTAATGCAAAACATTATCTGTAAATCAGAGATTACAGAAATTCATCGAG 840
QY 6464 GAGAGGAGCTGAATTTACAGAGATTACAGAGAAATTCCTTTTGGAGTCCGAG 6523
DB 841 GAGAGGAGCTGAATTTACAGAGATTACAGAGAAATTCCTTTTGGAGTCCGAG 900
QY 6524 TCGCAGCAGCCGGGGTAGAGAGCCGAGAGTGGCAGAAACCGTTCCGTAATCAATTTGCG 6583
DB 901 TCGCAGCAGCCGGGGTAGAGAGCCGAGAGTGGCAGAAACCGTTCCGTAATCAATTTGCG 960
QY 6584 GAAAGACGAGATGATGTGGAAAGAAAGCAAGATTTTGGTAAAAAGCTGATTAACG 6643
DB 961 GAAAGACGAGATGATGTGGAAAGAAAGCAAGATTTTGGTAAAAAGCTGATTAACG 1020
QY 6644 GAAAGACGATATGTTCTGTGGAAATGCGCAGTCCGAGTGGCAGAAAGAGG 6703
DB 1021 GAAAGACGATATGTTCTGTGGAAATGCGCAGTCCGAGTGGCAGAAAGAGG 1080
QY 6704 CTGGAATTTGAGCAGCAGATGCGAGTTACCAAGATGATCAACAGAGCAGAGTGA 6763
DB 1081 CTGGAATTTGAGCAGCAGATGCGAGTTACCAAGATGATCAACAGAGCAGAGTGA 1140
QY 6764 AAAATTTCTAAATTTATGACTCGAAACAGTTAGATTAATAGCAGAAATGAGATTAAT 6823
DB 1141 AAAATTTCTAAATTTATGACTCGAAACAGTTAGATTAATAGCAGAAATGAGATTAAT 1200
QY 6824 CAGGATCTGGAATCGGTTTACGCGAGCTGGAATTTCTTCCAGCAGAGATCTGAGAGT 6883
DB 1201 CAGGATCTGGAATCGGTTTACGCGAGCTGGAATTTCTTCCAGCAGAGATCTGAGAGT 1260
QY 6884 TTTCTGCTAATATATTTGCAAAATTAAGGTAGAAACAGATTCGATCACTTAATCACT 6943
DB 1261 TTTCTGCTAATATATTTGCAAAATTAAGGTAGAAACAGATTCGATCACTTAATCACT 1320
QY 6944 CTTCTAATGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 7003
DB 1321 CTTCTAATGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1380
QY 7004 GAGGAGCCGAGAGTCTTGGCAGCAGTACCGGAGTGGTTCTGTAAACATTAATATGTT 7063
DB 1381 GAGGAGCCGAGAGTCTTGGCAGCAGTACCGGAGTGGTTCTGTAAACATTAATATGTT 1440
QY 7064 CTGTAGTACTCGAGTTCAATTAATCTGATTTGACTCCGTAAGAGAAATGTAATG 7123
DB 1441 CTGTAGTACTCGAGTTCAATTAATCTGATTTGACTCCGTAAGAGAAATGTAATG 1500
QY 7124 TAAACGCAAAAGAGGAAAAATTAATTAAGCAACAGCAGCAATGAGCAATCGAGAG 7183
DB 1501 TAAACGCAAAAGAGGAAAAATTAATTAAGCAACAGCAGCAATGAGCAATCGAGAG 1560
QY 7184 CAGCAATCGAGAGCAATGTTCTGTGTAAATTAATTTGGAACAGCTGTAGAGAGTGA 7243
DB 1561 CAGCAATCGAGAGCAATGTTCTGTGTAAATTAATTTGGAACAGCTGTAGAGAGTGA 1620
QY 7244 ATTCTAAGGAAAAAGAGCAAGATTTTAAATCTTAAAGCAAGTAAACAAAGCAAG 7303
DB 1621 ATTCTAAGGAAAAAGAGCAAGATTTTAAATCTTAAAGCAAGTAAACAAAGCAAG 1680
QY 7304 ATTAATAAGTAAATGATGCTACGAAATAATCTTAAATCAGCAGGTAATTTCTACAGAG 7363
DB 1681 ATTAATAAGTAAATGATGCTACGAAATAATCTTAAATCAGCAGGTAATTTCTACAGAG 1740

QY 7364 ATACTCTGTAAACCGGATAGAGAGATCTCAGGGAGAAAGAAATTAAGCCATTGCA 7423
DB 1741 ATACTCTGTAAACCGGATAGAGAGATCTCAGGGAGAAAGAAATTAAGCCATTGCA 1800
QY 7424 AGACTCTGTATTAATTTGGAATAATGATATTTTAAACAGAGAGCAAGAAATATATCA 7483
DB 1801 AGACTCTGTATTAATTTGGAATAATGATATTTTAAACAGAGAGCAAGAAATATATCA 1860
QY 7484 CTTCTACTGAGGTTTGGGAACCTGAG 7510
DB 1861 CTTCTACTGAGGTTTGGGAACCTGAG 1887
RESULT 8
AR442021
LOCUS AR442021 1130 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 9 from patent US 6669940.
ACCESSION AR442021
VERSION AR442021.1 GI:42669172
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1130)
Negaraja, T.G., Stewart, G.C., Narayanan, S.K. and Chengappa, M.M.
TITLE Recombinant fusobacterium necrophorum leukotoxin vaccine and preparation thereof
JOURNAL Patent: US 6669940-A 9 30-DEC-2003;
Kansas University Research Foundation; Manhattan, KS
FEATURES
source
1..1130
location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 11.6%; Score 1130; DB 6; Length 1130;
Best Local Similarity 100.0%; Pred. No. 3.6e-119;
Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACCGGATCAAAATTAACGTTCAAGAGCAGAGAGAGATTCAGATTTCTAAAAA 60
DB 1 ATGACCGGATCAAAATTAACGTTCAAGAGCAGAGAGAGATTCAGATTTCTAAAAA 60
QY 61 GTTTAATGATTTGGGATGTTGATTAACCTATAGACGTGAGGCTTAATGATCAATC 120
DB 61 GTTTAATGATTTGGGATGTTGATTAACCTATAGACGTGAGGCTTAATGATCAATC 120
QY 121 ACCGGAATGAAATTTTGAACAAATAATAGAAAAAAGATTAATGATTAATCACT 180
DB 121 ACCGGAATGAAATTTTGAACAAATAATAGAAAAAAGATTAATGATTAATCACT 180
QY 181 ACAGCAAGATTAAGGAGAGACGCTTTTAAACGTTTAATAGATTTGCTTTAACAG 240
DB 181 ACAGCAAGATTAAGGAGAGACGCTTTTAAACGTTTAATAGATTTGCTTTAACAG 240
QY 181 ACAGCAAGATTAAGGAGAGACGCTTTTAAACGTTTAATAGATTTGCTTTAACAG 240
DB 181 ACAGCAAGATTAAGGAGAGACGCTTTTAAACGTTTAATAGATTTGCTTTAACAG 240
QY 241 AATATTAAGCAATCTATATTTTGGGAAAAAGATTAAGAGGAGTAAATATCTTTT 300
DB 241 AATATTAAGCAATCTATATTTTGGGAAAAAGATTAAGAGGAGTAAATATCTTTT 300
QY 241 AATATTAAGCAATCTATATTTTGGGAAAAAGATTAAGAGGAGTAAATATCTTTT 300
DB 241 AATATTAAGCAATCTATATTTTGGGAAAAAGATTAAGAGGAGTAAATATCTTTT 300
QY 301 AACTTTGTCATGAAAAATTAAGATAGATGATTAATCAACGAAATTCAGAAAAATTA 360
DB 301 AACTTTGTCATGAAAAATTAAGATAGATGATTAATCAACGAAATTCAGAAAAATTA 360
QY 361 ATTGAGAGAAATTAATTTCTTAAAGCTCGAAGGAGATGCGATGAGAAAAATGAGATT 420
DB 361 ATTGAGAGAAATTAATTTCTTAAAGCTCGAAGGAGATGCGATGAGAAAAATGAGATT 420
QY 421 ATCAATGCTGCTTTCTTCAATCTATTAATCCAAACAGAGATTTTAAGAGGCTTG 480
DB 421 ATCAATGCTGCTTTCTTCAATCTATTAATCCAAACAGAGATTTTAAGAGGCTTG 480
QY 481 GAAGAGCAACATGTTAAAGTTTAAAGATCAATTCAGATGAGAGAGTAAAT 540
DB 481 GAAGAGCAACATGTTAAAGTTTAAAGATCAATTCAGATGAGAGAGTAAAT 540

repeat_region	/rpt_family="ERV1"	11656	11747
repeat_region	/rpt_family="ERV1"	14141	14466
repeat_region	/rpt_family="MaLR"	15137	15194
repeat_region	/rpt_family="MaLR"	20668	20925
repeat_region	/rpt_family="B4"	21376	21458
repeat_region	/rpt_family="ERV1"	22395	22448
repeat_region	/rpt_family="Alu"	23930	24043
repeat_region	/rpt_family="MIR"	24193	24663
repeat_region	/rpt_family="Alu"	25105	25661
repeat_region	/rpt_family="Alu"	25263	26151
repeat_region	/rpt_family="ERVX"	26170	26253
repeat_region	/rpt_family="Alu"	26483	26644
repeat_region	/rpt_family="MIR"	27338	27756
repeat_region	/rpt_family="MaLR"	28247	28690
repeat_region	/rpt_family="MaLR"	29190	29265
repeat_region	/rpt_family="ID"	30206	30535
repeat_region	/rpt_family="ERVX"	30685	31038
repeat_region	/rpt_family="MaLR"	31361	31704
repeat_region	/rpt_family="Alu"	32067	32214
repeat_region	/rpt_family="Alu"	32231	32489
repeat_region	/rpt_family="L1"	32504	32715
repeat_region	/rpt_family="B2"	33729	34156
repeat_region	/rpt_family="L1"	34675	34935
repeat_region	/rpt_family="B4"	35007	35111
repeat_region	/rpt_family="L1"	36037	36186
repeat_region	/rpt_family="L1"	36325	36337
repeat_region	/rpt_family="Alu"	36828	36900
repeat_region	/rpt_family="Alu"	37058	37192
repeat_region	/rpt_family="Alu"	37595	37812
repeat_region	/rpt_family="B2"	37961	38151
repeat_region	/rpt_family="B2"	38477	38857
repeat_region	/rpt_family="ERV1	39224	39330
repeat_region	/rpt_family="B2"	39260	39576
repeat_region	/rpt_family="B4"	39823	39966
repeat_region	/rpt_family="Alu"	39981	40047
repeat_region	/rpt_family="B4"	40094	40221
repeat_region	/rpt_family="B4"		

```

repeat_region      41171. .41265
/rcp_family="B4"   43685. .43977
repeat_region      44014. .44224
/rcp_family="L1"   44245. .44485
repeat_region      44495. .44622
/rcp_family="L1"   44605. .46597
repeat_region      /note="Unresolved tandem repeat."
/rcp_family="L1"   46582. .47097
repeat_region      47292. .47735
/rcp_family="L1"   47832. .47886
repeat_region      /rcp_family="WALR"
/rcp_family="ERV1" 48245. .48397
repeat_region      /rcp_family="L1"
/rcp_family="B4"   50218. .50366
repeat_region      /rcp_family="B4"
/rcp_family="B4"   51688. .51899
repeat_region      /rcp_family="EMEREA"
/rcp_family="L1"   52035. .52749
repeat_region      /rcp_family="L1"

```

	Query Match	1.4%	Score 131.6;	DB 9;	Length 170221;
	Best Local Similarity	42.0%	Pred. No. 1,2e-08;		
	Matches 828;	Conservative	0;	Mismatches 1139;	Indels 3; Gaps 1;
Oy	6422 AGACAAAGATTGGAAAACTTACAATGCTCATGTAAGGTTTCGGAGAAGACTGAATTA				6481
Db	46596 AAMCAATGACAGTAGCAGTAGCAGAACGTCTAGCAGTAGCAGTAGCAGAAAGCACTACAG				4653
Oy	6482 CAGCAGATTCTAAGCAAGAAATTTCTCTTTTGAGTCGAGATCGCACGCCGGGTAG				6541
Db	46536 TAGCAGTAGCTGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGAAAGCTGTAGCAGTAGCAG				4647
Oy	6542 GAGCCGAGTAGGCAGGAACCGTTTCCGTAATATCAATTTGACGGAAACGGAGATAGATG				6601
Db	46476 TAGCAGTAGCAGAAAGCTGTAGCAGTAGCAGAAAGCTGTAGCAGTAGCAGTAGCAG				4641
Oy	6602 TGGAAAGACAAAAGATTTTGATAAAAAAGCTGAGATTACAGCAAACGTTATATGTTCTG				6661
Db	46616 TAGCAGTAGCAGAAAGCTGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAG				4635
Oy	6662 TTGCAATTGGAAATGCGCAGTCGAGTGGCTGCAAAAAGAGCTGGAATTGGACACAGC				6721
Db	46356 TAGCAGTAGCAGTAGCAGAAAGCTGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAG				4632
Oy	6722 TGCGAGTTTCCAAGATGAAATCAAAACAGAGAGCAABAGTGAATAATCTTAAATTTATGA				6781
Db	46396 TAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGAAAGCTGTAGCAGTAGCAGTAGCAG				4623
Oy	6782 CTCGAAACAAGTTAGATGTATAAGCAGAAAAATGATPAAATCAGGTACTGAAATCCGTT				6841
Db	46236 TAGCAGTAGCAGAAAGCTGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGAAAGCTGTAGCAG				4617
Oy	6842 CAGCCGAGCTGGAATTTCTTGACGCCGAGATCTGAGATGGTTTTCTGTCAATATATATG				6901
Db	46176 TAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGAGCAGCAGTAGCAGTAGCAGAAAGCTG				4611
Oy	6902 CAATTAAGATGAAACAGATATGCAATCACTACTTAACTTTTCAATGATGTAATG				6961
Db	46116 TAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAG				4605
Oy	6962 TAAAGCTCTTAATPAAATTTGCAATTCCTTGACAGCCGATGAGAGCCGAGGTCTTG				7021
Db	46056 TAGCAGTAGCAGAAAGCTGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAG				4599
Oy	7022 CAGCAGTTACCGAGATGGTTTCTGTAACTAATTAATAGTTCTGTGATAGCTGAGTTT				7081

[illegible][illegible]

RESULT 10	AC024253/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
	AC024253	179553 bp	DNA	linear	HTG_01-SEP-2000			
	Homo sapiens chromosome 6 clone RP11-758C19, WORKING DRAFT							
	SEQUENCE, 12 unordered pieces.							
	AC024253							
	AC024253.4	GI:9959959						
	HTG; HTGS_PHASE1; HTGS_DRAFT.							
	Homo sapiens (human)							
	Homo sapiens							

REFERENCE 1 (bases 1 to 179553)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 179553)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2000) Genome Sequencing Center, Washington
UNIVERSITY School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Sep 1, 2000 this sequence version replaced gi:8566958.

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0758c19
----- Summary Statistics -----
Sequencing vector: M13, 100%
Sequencing vector: plasmid, 0%
Chemistry: Dye-primer ET, 100% of reads
Chemistry: Dye-terminator Big Dye, 0% of reads
Assembly program: Phrap; version 0.99019
Consensus quality: 175235 bases at least Q40
Consensus quality: 176522 bases at least Q30
Consensus quality: 177171 bases at least Q20
Insert size: 198000; agarose-fp
Insert size: 178453; sum-of-contigs
Quality coverage: 4.48 in Q20 bases; agarose-fp
Quality coverage: 4.99 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
7554: contig of 7554 bp in length
7555 gap of unknown length

```


QY 6175 ACTTTAAATCTTTATTGGCAAAATCCGCTGGTTCAGACAAAGCCGAGTGGCAGAACT 6234
Db 175564 GCACACGACACAGCAGAGTAGTACCAACAGAAATCGCATGTAATTAACAGTAGCAGACAGCA 175505
QY 6235 GTTAATATCAACAAGTTTATGAGAAACAGAAAGCTCTTGTGAAATTTCTATATTAAT 6294
Db 175504 GAGACATATTTACATAGTAGCAACACAGTAGCAGCAACCCAGCAGCAATATATAGCA 175445
QY 6235 GCAAAACATTTATCTGTAAATTCAGAGATTACAGAAATTCAGATCGAGTAGTGTCT 6354
Db 175444 GCAATAGCAACAGTAGATATAGTAGCAACAGCAATAGCAACAGTAGCAGCATACAGACA 175385
QY 6355 GTTGTGTTGTGAGAAATGAGAGTAGAGGCTTCTTGTATACCAATTTATTAAGA 6414
Db 175384 ACAGCAACAGCAGCAGCAACATTAAGAGCAGCAGCAACAAACAGTAGCAGACAGCA 175325
QY 6415 AATACCAAGACAAGAGTTGGAAAACTACATGCTGATGAAAGTTTCGGAATAAGAT 6474
Db 175324 GCACACACGACACATTAACAAAGACAGATAGCAACAGTATCAGCAACAGCAACAGCA 175265
QY 6475 GAAATTTACAGCAGATTCTAAGCAGAAATTTCTCTTTTGGAGTCGAGATCGCAGCAGC 6534
Db 175264 TCAGCAACACAGCAACATATACAGCAGTAAAAACAGCAACAGCAGCAGCAACAGCA 175205
QY 6535 GGGGTAGAGCCGAGAGTGCAGAAACCGTTCCGTAAATCAATTTGCAAGAAAGACGGA 6594
Db 175204 ACAGCAGCAGCAACATTAAGACAGCAACACAGCAACAGTAGCAGCAGCAACAAACA 175145
QY 6595 GTTAGATGTGGAAGAAAGCAAGTTTTGTAAAAAAGTGAATTAACAGCAAAAGTTAT 6654
Db 175144 GCACACACACAAAGCAACAAATAGCAACAGTATCAGCAACAGCAACAGCAATACCAACA 175085
QY 6655 AGTTCTGTTCGAATTTGAAATGCCGAGTCGAGAGTGGCTGCAAAAAGAGC--TGGAAAT 6711
Db 175084 AAGAGATCAGCAACAGCATATAGCATCAATTAACACAGCAACAGCAGCATATACACA 175025
QY 6712 GAGCAGCAGTGGCAGTTACCAAGATGAATCAACAGAGCAAGAGTGAATAATCT 6771
Db 175024 GAAATCAGTAGTAACCAAGTAGCAGCAGCAGCAATTAATTCATATAGTACCAACA 174965
QY 6772 AAAATATATGCTCGAAACAAAGTTAGTGTAAATAGCAAAAATGAATTAATCAGATCT 6831
Db 174964 GTAGCAGCAACACAGCAGCAGTAGAGAGCAATAGCAACAGTAATATAGTACAGCAACA 174905
QY 6832 GGAATCGTTTCAGCCGAGCTGGAATTTCTTGACCCGAGATCTGGAATGTTCTGTC 6891
Db 174904 GCAATAGCAACAGTAGCAGCATCAGCAACAAAGCAACATAGCAGCATTAACA 174845
QY 6892 AATTAATTTGCAATTAAGTAGAAGATATGATCATAGTACTTTACATCTTCTACT 6951
Db 174844 GCACACAGCAGCAACATTAAGAGCAGCAGCAACAAACAGTAGCAGCAGCA 174785
QY 6952 GATGTAAATGTAAAGCTCTTAATAAATTTTGAAATTCCTTGACCCGCTGGAAGAGC 7011
Db 174784 ACACACGCAACATTAACAAAGCAACAGTAGCAACAGTATCAGCAACAGCAACAGCAATA 174725
QY 7012 GCGAGCTTTCAGCAGTACCGAGTGTCTGTGTTACATATAATAGTCTGTGATA 7071
Db 174724 GCACACGACAGCATCAGCAACATAGCAGCAGTAAAAACAGCAACAGCAGCAGCAACA 174665
QY 7072 GCTCGAGTTCACAATTACTGTGATTTGACTTCGTGAGAGAAAAAGTAATATTAACGGA 7131
Db 174664 ACAGCACAACAGCAGCAGCAACATTAAGAGCAGCAACAGCAACAGTAGCAGCAGCA 174605
QY 7132 AAAAGGAAAAAATTTTAAGCAACAGCAGCAATAGCAGAAATCGAGAGAGCAGCAATC 7191
Db 174604 ACACACGCAACATTAACAAAGCAACAGTAGCAACAGTATCAGCAACAGCAACAGCAATA 174545
QY 7192 GAGGCAATGTCTTGTGTAATATTTTGGAAACAGCTGTGAAGATGAATAAATTTCTGAA 7251
Db 174544 GCACACAAAGCATCAGCAACAGCAATAGCATTAATTAACACAGCAACATATAGCACA 174485
QY 7252 GGAAGAGAACAGAGTTTAAAACTTTAGAGGAAGTTTAAAGAAACAGATTAATAA 7311

Db 174484 GCACACAGCAACAGCAGTAGTACCAACAAATTCAGTAATTAACAGTAGCAGCAGCA 174425
QY 7312 GTAATGATGCTAGAAAAAATCTTAACATCAGAGGATATTTCTACAGAGATCTCT 7371
Db 174424 GCATCATTTACATATGTAAGCAACACAGTAGCAGCAACACAGCAGCATATGACAGCA 174365
QY 7372 GTAAAGCGGATTAAGAGATCTCAGGAGAAAGAAATTAAGCCATTTGTGAAGCTTCT 7431
Db 174364 ATAGCAACAGTAATATGTAAGCAACACAGCAATAGCAACAGTAGCATCAGCAACACA 174305
QY 7432 GATATTTGGAAGAAAAATGTAATTTTCAACAGAGCAAGAAATATATCATCTTCTACT 7491
Db 174304 GCACACGACAGCAAC-----AATAGACAGCCGAGCAACAGCAACAGTAGCAGCA 174251
QY 7492 GGTGTTTGGGACATGACAGTCTTGCTTCCGATCAGAGCAAGTGGCAGTTACAAATAT 7551
Db 174250 GCACACACACACAGTATGACAGCAGCAGCAACACAGCAACATTAACAAAGCAACATA 174191
QY 7552 AAAAGAAATTCGAGATTAATCTGTGAAATTTCTTTGAAAGCAGCTGAAAAATTAAT 7611
Db 174190 GCAACAGTATCAGCAACAGCAACATAGCAGCATCAGCAACAGCAACATATAGCACA 174131
QY 7612 GTTAGATGATATTTACAGAAATGTTGCTTTAACAGCATATCAAGTCTGTAGAGCA 7671
Db 174130 GTAAAAACAGCAACAGCAGCAGCAACACAGCAACAGCAGCAACATTAAGAGCAGCA 174071
QY 7672 TTGGGAATAGAGAGCTGCTTACAGAAATTAATTTAATGGAAGTCAATATCAGTAT 7731
Db 174070 ACACACGCAACAGTATGACAGCAGCAGCAACAGCAACATTAACAAAGCAACATAGCAACA 174011
QY 7732 AAAATTTAAGCTATTAAGAAAAAATTTAGATTTATTTGAAAGATMAATCGGAATTC 7791
Db 174010 GTATCAGCAACAGCAACAGCAATTAAGAAACAAAGATCAGCAACAGCAATAGCATCAACA 173951
QY 7792 AGAGCGAAGCAAAAGATTAACGTAAGAGCGTATGCTGCCGAGCATTAATCTCAAAA 7851
Db 173950 TTAACACAGCAACATGAGCCAGCAGCAACAGCAACAGCAGCAGTATAGCAACAAATCA 173891
QY 7852 GCAAGAAATGAATGAATTTAGAGAGTTGAATTTGAAGAGATTTTCAATGAAGAAAT 7911
Db 173890 GTAAATTAAGTAGCAGCAGCAGCAGCAGCAGCAGCAATTCATTCATATGTAAGCAACACA 173831
QY 7912 AGAGTAATAGCCCTTTTAAAGAAATTTGGAAGAAATCAATGTCAAAAGTGAAGAA 7971
Db 173830 GTAGCAGCAACACAGCAGCATATAGTACAGCAATAGCAACAGTATATGTAAGCAGCAACA 173771
QY 7972 AACAGAGTACTGCTGAATCTCAG-----GAGCTTCTGTAGAGCAGTAGAGGGCA 8025
Db 173770 GCAATAGCAACAGTAGCAGCATTCGCAACAAAGAGCAACAAATAGCAGGTAAACAACAGCA 173711
QY 8026 GGAATTAATTTCCGAAGCAAAAGATGCCGGAAGCTTTATTTGAAGTTAGTACAAATCC 8085
Db 173710 ACAGCAGCAACATATGAGAGCCCGCAGCAACAGCAACAGTAGCAGCAGCAACA 173651
QY 8086 GGAAGAAATATTTTTCATGACAGATATGTGATATGAAAGCAACATTAATGAAGTA 8145
Db 173650 ACAGCAACATTAACAAAGCAACAGTACCAACAGTATCAGCAACAGCAACAGCAATAGCA 173591
QY 8146 ACAGCATTTCTAAAGCAGTAACAGTCTGTATTTGGAGGAGTTGGAGTCAACAGCA 8205
Db 173590 ACAGCAGCATCAGCAACAGCAATAGCATCAACATTAACACAGCAACATTTAGCCACAGCA 173531
QY 8206 GAAGCTACTGCTGCAAGTAAC-----TATGTAGAAAGTTGAGGAAGAAAT 8253
Db 173530 ACAGCAACAGCAGCAGTAGCAACAGAAATAGCAGTAATAGCAGCAGCAGTAGCAGCA 173471
QY 8254 TTGTTCAAGCAACATTCAGATTGAATGCAATTTCTAAAGTAGAAGGTTGATGAAGATTA 8313
Db 173470 AATAATTAACATATAGCAGCAATAGCAGCAACACAGCAGTATAGAGCAGCAACAAATTA 173411
QY 8314 GTAACTGTAATCTTCTGTAGTATCAGAAATGAGAGGAATTCGAGAGCAGAGTG 8373

Db 173410 AAGCAACAACAGCAATAGCAGCAATCATTAACAGCAATAGCAGCAGCAACA 173351
Qy 8374 AATACCTTACGCGCAAGTAATATCTGAATCCGTAAGTCTTACGAAAGCAAGTAT 8433
Db 173350 ACGCAGTAACAAATAGCAGTAATGACAACTACGACGACGACCAACAGTGTCA 173291
Qy 8434 GAAATTAATGATTAACCAAAAAAATATATTGAGATCATGCTCTTGAATGAT 8493
Db 173290 GCAACGTAAGCAGCAGTACAGCAATAGCAACGATCTATTAACAAATGCGCAATTA 173231
Qy 8494 ACAAAGTAAGACGAAATTAAGATCTTAAAGCGTAGCCGATGTCAGTACAGCAAGACA 8553
Db 173230 ACAGAGCAGCAGCAACAAACAAAGCAACATAGCAATAGTACAGCAACCAACA 173171
Qy 8554 AACAAAGCTTTTACGAGATCAACAAAGTTACTTACACACTGTAATGAGAAACGTA 8613
Db 173170 GTACAGTACGACGACCACTACAGATGTAACAAACGAGTTCGAGTAATAGTGA 173111
Qy 8614 TCTCAACTTCGTAACAAAGCTTTGGCTAAATAATGAAATGTAAGAAAGAACT 8673
Db 173110 ACAACGACGACGACGACCACTAGTATGACAAATGACACGACTTACAAACACGACACA 173051
Qy 8674 GGAGAGCCTTATAGTCGAGCGGAAACAGACGCCGTTGAAATTTATACAA 8722
Db 173050 ATTTACGACACGACGACGACGACGACGACGACGACGACGATCGTATGACAA 173002

RESULT 11
PFAMESA 4995 bp mRNA linear INV 26-APR-1993
LOCUS Plasmodium falciparum (MESA) mRNA exons 1-2, complete cds.
DEFINITION
ACCESSION M69183
VERSION 69183.1 GI:160408
KEYWORDS erythrocyte surface antigen.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 4995)
REFERENCE Coppel, R.L.
AUTHORS Repeat structures in a Plasmodium falciparum protein (MESA) that binds human erythrocyte protein 4.1
Mol. Biochem. Parasitol. 50 (2), 335-347 (1992)
JOURNAL 1741020
PUBMED Original source text: Plasmodium falciparum (strain D10) cDNA to mRNA.

FEATURES
source
1..4995
/location/Qualifiers
/organism="Plasmodium falciparum"
/mol_type="mRNA"
/strain="D10"
/db_xref="taxon:5833"
/join(278..379,502..4932)
/gene="MESA"
/join(278..379,502..4932)
/gene="MESA"
/codon_start=1
/product="mature-parasite-infected erythrocyte surface antigen"
/protein_id="AA29651.1"
/db_xref="GI:160408"
/translation="MENBGNKVKVYNNSSLKRYMFCCTIICVFLIDITYNCEPT
YYSISIKNNNDYRVRILSETPPMSLIEIRTPDEHLISIRNYICLNAPYIDPL
WGSVVDKRNKCLQHIKLEMOSESRKQOEBNAKDIIEIKKEKYLKLEMD
ESDVEKAFRELOI KLRDTRPKHVVNMGESKETDESKEITGESKETGESKE
TGESKETGESKETGESKETGESKETGESKETGESKETGESKETGESKETGES
KETGESKETGESKETGESKETGESKETGESKETGESKETGESKETGESKET
ESKORGNKYSQPYENBNSNVTSESSETKCLAEKNEBERKUGENVNGASNSDP
KTLTEBENGTKSESSETKODRPENEKADNKKKRRKKKSFQMLGKCNFLCKNTE
TDEBETLVVKDIAKKHKKFLREANTEKDNKKDLKLGDSKEDVKEKNDKQKVL
GGEDKDVKEKNDKQKVLGGEDKEDVKEKNDKQKVLGSEKTKKEIKYKVKV
KKCKKVKKGIKENDEKNDKVKPEIIEVKEEIKKQVEDGIKENDEKNDKVKP
EITTEVKEIKKQVEDGIKENDEKNDKVKPEIITTEVKEIKKQVEDGIKENDE
GNDKVKPEIITTEVKEIKKQVEDGKENDTESKVKIGOEIITTEVKEIKQEBK

GNKENLIEIKDVIIGOEVIIEVKVKIKKVEKGIKENHTESKDKVIGOEIIEVEKE
EIEKOVEEGIKENDTESKDKVIGOEVIIGADVNESGPEPKDKVYKQEKVEKVEK
VKKVKRNKNNKNDVYIGKIMKEDVNEOTANKDKEIOEKEKEKVEKVEK
EVEKVEKVEKVEKVEKVEKVEKVEKVEKVEKVEKVEKVEKVEKVEKVEK
KDKVIGOEIIEIIEIKVEKVKRNKNNKNDVYIGVINEKEDVNEOTANKDVI
OKEKEVEKEVEKEVEKEVEKEVEKEVEKEVEKEVEKEVEKEVEKEVEKE
VNEKTESKDKMIGKEVIIEBEKVEKVKVEKVKVNRNRRKNERKDOVIEQIVSE
EVEKEDTGNKDKIGKRVKCPIDCKKEREVOIINVEDNEKTESKDKMIGKEVI
EVEKVKRVKVEKVKVNRNRRKNERKDOVIEQIVSEVEKEDTGNKDKIGKRVK
PIDCKKEREVOEBSSESESESESESESESESESESESESESESESESESESE
SE
ONEVFEDEYIIQSKENHTLNTFPNMALENDFRFRFPHNLIETHDTLHMKRIIYNL
FLEVFPNNKOKAKELDKKFSBMNVHQLNLYBERIRYEBENAAGRNLTVILDPIT
FNVIIGDDTWFKFIDE"
278..379
/gene="MESA"
/number=1
380..501
/gene="MESA"
/number=1
502..4932
/gene="MESA"
/number=2

ORIGIN

Query Match 1.3%; Score 123; DB 2; Length 4995;

Best Local Similarity 43.6%; Pred. No. 6.3e-07; Matches 1032; Conservative 0; Mismatches 1290; Indels 45; Gaps 9;

Qy 7110 AGAAAAAGTAATGTACGCGCAAAAAGCAAAAAATATTAAACAAACGACAAATGC 7169
Db 1482 AGAGACCAAAAATTAAGCGGAAAAAGAGATGCGGAGAAAAATTAAGGAAAAATG 1541
Qy 7110 AGCAATCGAGAGACGACCAATCGAGCAATGCTTTGTAATTAATTTGGAACAGCTGT 7229
Db 1542 TAATGATGGCGCATAGAAAATTCAGAGATCCC-----AAAAATTAACGAAACMAA 1595
Qy 7230 AGAAGTGAAAAAATTCGAAAGAAAGGAAACAGAGTTTAAAACTTTAGACGAAGT 7289
Db 1596 AGAAAAATGTACAAAAGAAAGTTCTGACAAAACAAAGATGATTAACCGAAAGAAATGA 1655
Qy 7290 TAACAAAGAACAGATTAATAAGTATGATCTACGAAAAAATCTTCAATCAGCAG 7349
Db 1656 GAAAAAGCAGATTAATTAATAAGTAAAAAAGAAAAAATATCTTTTCAATGT 1715
Qy 7350 TAATTTCAAGAGATCTTCTGTAAGCGGATAGAGAGATCTCAGGAGAGAAAT 7409
Db 1716 AGAGTGAATTTCTTAATTAATAAATAATTTGA---ACCTGATGAGAAAGAAACGTT 1772
Qy 7410 TAAAGCATTTGGAAGATCTTCGATATTATTGGAATAAATGTATTAACAACAGAGA 7469
Db 1773 GGTAGTAAGAATGATGCTTAAGAAAGAAACATTAATTTTAAGGAAGCTTAATCTGAAA 1832
Qy 7470 CAAGATAATATCATCTTACTGATGTTGGAACTGACAGGCTTGTCCGATCAG 7529
Db 1833 AATGATTAATGAAGAAGAAATTAATTAATTAAGGAAGAGATTAAGAGATTTAAAGA 1892
Qy 7530 AACAGTGCAGTTACAAATATTAAAGAAATTCGAGATTACTGTTGAATTCCTTTGT 7589
Db 1893 AAAGATGATGAACAGAAAGATTAAGTATTAGAGAAAGAGATTAAGAAAGATGTTAAAGA 1952
Qy 7590 GAAAGCAGTGAAGAAAGTAAATCTTGAATCGGATATTACAGAAAGTTGCTTAAACAG 7649
Db 1953 AAAGATGATGAACAGAAAGATTAAGTATTAGAGAAAGAGATTAAGAAAGATTTAAAGA 2012
Qy 7650 ATATCAAGTCTCTGAGAGCAATGGGAATAGAGCTGCCTATGCAAGATTAAATTTAA 7709
Db 2013 AAAGATGATGAAGAAAGATTAAGTATTAGATCAAGAAAAACCAAGAAAGAAATTA 2072
Qy 7710 TGAAGATCAAAATATCAGTATTAAAAATTTAAAGCTATTAGGAAAAAATTTGATGTTAT 7769
Db 2073 AGAAAAAGTGAAGAAAAAGATTTAAAA---AAAGTGAAGAAAAAGTAAAAAAGAAAT 2129

Medical Center Dr., Rockville, MD 20850, USA

FEATURES

source

1. 271546
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="11"

repeat_region
128..152
/rpt_type=tandem
complement(130..207)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
150..183
/rpt_type=tandem
complement(313..373)
/rpt_type=tandem
/rpt_unit="(a)n"

repeat_region
313..343
/rpt_type=tandem
complement(426..445)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
complement(485..504)
/rpt_type=tandem
/rpt_unit="(a)n"

repeat_region
complement(644..683)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
644..671
/rpt_type=tandem
complement(786..836)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
799..836
/rpt_type=tandem
complement(982..1091)
/note="AT rich"

repeat_region
/rpt_type=tandem
complement(1002..1024)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
1148..1175
/rpt_type=tandem
complement(1148..1175)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
complement(1245..1286)
/rpt_type=tandem
/rpt_unit="(a)n"

repeat_region
complement(1346..1412)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
1346..1373
/rpt_type=tandem
complement(1380..1446)
/rpt_type=tandem
1384..1433
/rpt_type=tandem
complement(1412..1446)
/rpt_type=tandem
/rpt_unit="(a)n"

gene
1534..2550
/locus_tag="PF11_0457"

mRNA
1534..2550
/locus_tag="PF11_0457"

CDS
1534..2550
/locus_tag="PF11_0457"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA036037.1"
/db_xref="GI:23496381"
/translation="MEKEKITYDRQLRMGVKAKQNRMKSNVLYVGLSGINIELCKNL"

LANGINITIIDNNIVDEEDIEINFFLNHEMKYNSVPIFKELKAINOLINIKSFIGHT
NLDITNNIITIEKELMYKDNIEITOEKRIINDCSYDIDCYITNTVCVISCEDYPLKLT
INIEIChENKIGFFANMGKAYAFLEPFGKHIIIESYKIKNDONKGNDSIEVOYC
TISHPLKVPESNLDKTNDIIYYFALILYEODKKMKONKOIDEQELNFTNNKLANK
TNTONPFIELCKTYIKNSPSCSINGCVTSQEIRKFIQRHESINFCVFDNNQIVCT
SMIK"

repeat_region
complement(1776..1830)
/note="AT rich"

repeat_region
/rpt_type=tandem
2544..2577
/rpt_type=tandem
complement(2555..2620)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
2575..2620
/rpt_type=tandem
complement(2664..2702)
/note="AT rich"

repeat_region
2720..2748
/rpt_type=tandem
/rpt_unit="(a)n"

repeat_region
2750..2775
/rpt_type=tandem
complement(2750..2775)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
2843..2881
/rpt_type=tandem
complement(2843..2881)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
complement(2929..2962)
/note="AT rich"

repeat_region
2990..3019
/rpt_type=tandem
complement(2990..3019)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
complement(3027..3056)
/note="AT rich"

repeat_region
complement(3113..3137)
/note="AT rich"

repeat_region
complement(3160..3182)
/note="AT rich"

repeat_region
complement(3221..3263)
/note="AT rich"

repeat_region
3350..3401
/rpt_type=tandem
complement(3350..3401)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
complement(3414..3506)
/note="AT rich"

repeat_region
3503..3555
/rpt_type=tandem
complement(3507..3576)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
complement(3578..3601)
/note="AT rich"

repeat_region
complement(3667..3748)
/rpt_type=tandem
/rpt_unit="(ta)n"

repeat_region
3714..3748
/rpt_type=tandem

```
repeat_region complement(3760..3796)
                /note="AT_rich"
                /rpt_type=tandem
repeat_region complement(3873..3924)
                /rpt_type=tandem
                /rpt_unit="(tca)n"
repeat_region complement(3942..3967)
                /note="AT_rich"
                /rpt_type=tandem
                complement(4068..4164)
                /note="AT_rich"
                /rpt_type=tandem
                complement(4087..4119)
                /rpt_type=tandem
                /rpt_unit="(taaa)n"
                <4155..>5108
                /locus_tag="PF11_0458"
                <4155..>5108
                /locus_tag="PF11_0458"
                4155..5108
                /locus_tag="PF11_0458"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AA036038.1"
                /db_xref="GI:23496382"
                /translation="MNSKENDSVSKLSEKNNIKLIVKPEVSNKMLQYNKDIVGL
                LQNNNNNSRITELYQKDDSNVKKLRCKNNTVTNYIILKQVYIKLSTNKSIVSN
```

Query Match 1.2%; Score 121; DB 2; Length 271546;

Best Local Similarity 42.8%; Pred. No. 2.6e-07; Matches 1766; Conservative 0; Mismatches 2265; Indels 96; Gaps 19;

```
QY 5120 ATCGTGGTCTGATGAAACGGAAGCTTTAGTATTCTGAGTTTGAAGAGTAAAGTT 5179
Db 188710 ATACTGAATTTGTTGTAAGAAAGATCAAGCTACTGAATTTGTTGTAAGAAAGATCG 188769
QY 5180 CTTTCATGTAGTGAAGAAAGATCAAAAAACAATAATCATTTGCCGGAATG---CAA 5236
Db 188770 CTACTGAAGTAGTTGTGAAGAAAGATCAGCTACTGAAGAGTTGTGAAGAAAGATCAG 188829
QY 5237 ATGGAAGAAAGAGCGCTGAGTTGAGCAACAGTTGCTCATCAATATTGAAAAACAT 5296
Db 188830 TTTCTGAAGAAATGTTAGAAAGAGATCAGCTACTGAAGAGTTGTGAAGAAAGAT 188889
QY 5297 CAGTTATAGCTATTGTAAAAAACAATAATTACACGCGAATGATCAAGATAGAAAA 5356
Db 188890 TATCTAGTATATGTTCAAAAGTCAAAAGAGTATTGAAAACGTAGAGAAATATATT 188949
QY 5357 ATATCAATGTACTGCAAAAAGATTATATGACCAATACATACTATAGCAATCGGAGTTGGAG 5416
Db 188950 CTGTAAAAACAGCTAAGATGATCTATGATGAAAAAATACCATTAGAGAAATCATCAT 189009
QY 5417 GAGCAAAAGAGAGCTGTGCAAGAGCTTCTGCAAGTACTACTTGAATTAAGACGTTT 5476
Db 189010 TTGTTGATATAGTCTCTTTAAAGACAAAGACCTACAGATATATGATGCGAAGAGC 189069
QY 5477 CTTCATGTTGTTCAAACTGATATATGACAAAGATTAGAGAAAGAAATTAATGAAATA 5536
Db 189070 TAAATTCGGAAGATATTTAATTAAGACATCTCCATTAGAGAAAGAAATTTGAAAGAT 189129
QY 5537 AGGAAAAAGCAATGTTAAATGTTCTAGCTGAAATAACA--GTCAAGTGTCACAAAATG 5593
Db 189130 TACCAACAGAAATATATTACTACTGCTGATATACACACTAAAGGTGAAAACAGAAACAAAT 189189
QY 5594 CGACAGTGTCTTCCGAGAGCAAGTGAACAGCTGCATAGAGCTGAGAGTGAAGCATTTAA 5653
Db 189190 ATATATCTAATATGCAAAAAATTAATGAAGAGTTGAAAAAGCAAAATTTCAAGAGAAA 189249
QY 5654 AAATTACACAAATACTCTGCACATATAAAAAATAGTACTCAAAATGTACGAATGCTT 5713
Db 189250 AAATTTCGGAATATTTCCAGTGAAGAAAGAAATCAAGTTACAGAGAGTATATAGTCAAGAAC 189309
QY 5714 TGGTAAAAAGCAAACTCATTCATCTATTAAAACAATTGGAATGAGCTGAGTTGGAG 5773
```

```
Db 189310 CATCATTTAGCTCAGGAAGTTGAACAAAAAGATCAGATACTAATGAATTTGAA---GAAA 189366
QY 5774 CTGAGAGAGCTGAGTGAACAGGTTCTGTAGCAGTGAATAAGATTGTAATATATACATAG 5833
Db 189367 CTAAATTGCTTAATGAAGAAAAATTTATACAGAGATATCTGTTACTGAAAAAGTTGTTGAAA 189426
QY 5834 CAGAAATTAATCATGCAAAAAATCACTGCGAAAGGAATGTGCGAGTTATTACAGAGCTG 5893
Db 189427 AAGAAAGATTAAGATCTGAAAGAGGCTTATGAAGAAAGATGAATCAATTACTGAAAGAAATAG 189486
QY 5894 ATGCGTAATTTGCTAATTTATGACAGAACAGTGTCTGAGTGGCCGTCAGCAATAGGAG 5953
Db 189487 TAGAAGAAAGATATCATCTAGTGAAGAAATAGTAGAAGAAAGAAATCATCTAATGTAAG 189546
QY 5954 CCTCAACCAAGTGAATGAATTAACAGAGCTTACAAAGCATATGTATAAAGATTCTACAG 6013
Db 189547 AAATGTAGAGAGAAAGAAATCATTAAGTGAAGAAATGTAGAGAAAGAAATCATCTA 189606
QY 6014 TGATTTGCTAAAGAAAGAAACAGATGATTAATTAATCT-----ACTCAAGGCGAAGTAGATA 6067
Db 189607 GTGAAGAAATATGTAGAGAGAAAGAAATCATCTAGTGAAGAAATATAGTAGAGAGAAAGAT 189666
QY 6068 AAGTGTATGATTAAGTATTCAAAAAATCTTAATTTATTAAGAAAGCTTATCAAAAAAGAA 6127
Db 189667 CATCTAGTGAAGAAATTAATGAAGAAAGTCAAGTACTGAAGAAAGCTTTAGAAAGAAAG 189726
QY 6128 AAATAGTATATAAAAAAGATTTGTTACCAATAGTTGAGTCACTACTCATCTTTAAATCTT 6187
Db 189727 GATCGTTTACGAGAAATTTGTAGAGAAAGATATCACTACTGAAAGAAAGTTAAAGTA 189786
QY 6188 TATTGCAAAATGCCGCTGTTTCAAGCAAGCCGAGTGGCAAGAACTGTAATATCAACA 6247
Db 189787 TTGATTCAGTTAGCGAAGAGATTAAGAAAGAAAGATTCGGTACTGAAAGATTTGTAG 189846
QY 6248 AAGTTTATGGA---GAACAGAAAGCTCTTTGAGAAATTTCTATATTAATGCAAAACATT 6304
Db 189847 GACAAAGAGATCTGTTATTGAAGAAATTTGAAAGAAAGAAATCACTCTGAAAAAGTTTC 189906
QY 6305 ATTCTGTAATAATCAAGAGATTAACAGAAATTCATCGAGTATAGTCTGTGCTGTTG 6364
Db 189907 TAGAAGATTTGATCAATTTGTGAAGAAATAGTACAGAAAGAAAGTCTGTGCTCAAG 189966
QY 6365 G---TGAATGTAGAGTAGAGAGCTTCTGTATACCAATATTAAAAAGAAATACCA 6421
Db 189967 AAATTTGATCAGAAAGATATCACTACTGAAAAAATGATGATGATGATCAACTTA 190026
QY 6422 AACAAGATGGAAGAACTA-----CAATGTCTGATGAAGTTTCGAGAAAGAGCTG 6475
Db 190027 CTGAAGAAATTTGTGAAGAAAGTTGATCAGTTCTGAAGAAATTAATAGTGAAGAAAGTGT 190086
QY 6476 AAATTAAGCAGATCTTAAGCAAGAAATTTCCCTTTTGGAGTGGAGTCGACAGCCG 6535
Db 190087 CAGCTTCTGAAGAAATAGTGAAGAAAGATCAGTTACTGAAAGAAAGTTGAAAGAGAAA 190146
QY 6536 GGGTAGAGCCGAGTAGCAGAAACGTTTCCGTAAATCAATTTGACAGAAAGACGAG 6595
Db 190147 AATTAAATTAAGAGTTGTGTAACCGAATCGGTTACTGAAAGAAATAGTACAAAAAGAA 190206
QY 6596 TAGATGTGAAGAGCAAGATTTTGTGTA-----AAAGTGAATTAACAGCA 6646
Db 190207 TATCAGATGCTGAAGAAAGTTTATAGCAAGAAAGATCAATGATTAAGAAATATTAGAA 190266
QY 6647 AACGTTATAGTTCTGTGCAATTTGAATATGCCAGCTGCGAGTGGCTCAAAAGAGCTG 6706
Db 190267 AAGAAATCAATTTGTTAAGAAATTTGTTGACCGAAGAGATCTGTTACTGAAGAAATAGTAG 190326
QY 6707 GAATTGAGACA---GCAATGCAAGTTACCAAGATGAATCAAAACAGAGACAGAGTGA 6763
Db 190327 ATCAGAGATCAATTTGCTGAAGAAAGTCAAAAGAAAGAAATTTAGTTACAGAAAGAGCTGTAC 190386
QY 6764 AAATTTCTAAATTAATGACCTGAAACAAAGTTAGATGTATATGACAGAAATGAGATTAAT 6823
```

Db 190387 AGTATGAAGGATCTGTACTGAAAGAAATATAAGAAAGAAATGCATTAAGTACGAAATGAAG 190446
Qy 6824 CAGTACTGGAATCGGTTTCAGCCGAGCTGGAATCTTGCAAGCCGAGTATTCGAAGTGG 6883
Db 190447 CAATGAAAGAAATCAGCTTTTGCAAAATATATGAAAGAAAGAACCAATATCTGACGAA 190506
Qy 6884 TTTCTGCTCAATATTTGCAAAATAGTAGAAACGATATCGATCATGACTTACTTACACT 6943
Db 190507 TTGTAAAGAAAGAAAGATTAAGTACTGAAAGAAATGTGATAGTATGATCACTACTGACG 190566
Qy 6944 CTTCTACTGATGTAATGTAAAGCTCTTAATAAATTTGCAATTCCTTGACAGCCGGT 7003
Db 190567 AGGTATGAGAAAGAAATTAAGTTAATGAACAATTTAGTAGAAGAAAGATCTGTTA 190626
Qy 7004 GAGGAGCCGAGCTCTTGACAGATTTACCGAGTGGTTTCTGTAACTATTAATAGTT 7063
Db 190627 CTGAAAGAACAGTAGAAGTTGAACATCAGTACTGAGATCTGTGTAAGAAAGACAT 190686
Qy 7064 CTGTGATAGCTGAGATGCAATTAATCTGATTTGACTTCGATGAGAAAGATPAATG 7123
Db 190687 CGGTATCTGAAAGAGTTTCACTGATGAAGATCAACTAATGAC--AAATGCTAGATG 190743
Qy 7124 TAAAGCGCAAAAGAGAAAAAATATTAAGCAAAACAGCAAAATGCAAGATCGAGAG 7183
Db 190744 AATCAGTTGCTGAAGAAATAGTAGAGAAAGAGTATCTGTGATGCAAAATTAATAGAG 190803
Qy 7184 CAGCAATCGAGCCCAATGCTTGTGAATATTTTGGAAACAGCTGTAAGATGAGAAAA 7243
Db 190804 AAGAGTATCTGTGATGAAGTTGTGAAGAAAGATCAGTTATTTGAAGATAGTAG 190863
Qy 7244 ATTGGAAGAAAGAAAGAAAGTTTAAAACTTTAGACGAATTAACAAAGAAAGCAAG 7303
Db 190864 AAGAGAAACAAATCAGTTCTGTGAAGAAATTTGGAAAGAACTATCAGGTTCTGAAGAG 190923
Qy 7304 ATAAAAAGATTAATGATGCTAGCAAAAAAATCTTAATCAGCAGGTAATTTCTACAGAG 7363
Db 190924 TTCTAGAAACGATAGGAGTTACTGATGCTTTTATGGAACAAGAGATCTGTTTAAAG 190983
Qy 7364 ATACTTCTGTAAAGCGGATGAGAGATATCTCAGGAGAAAGAAATTAAGCCATGTGA 7423
Db 190984 AATATGAAGAAATAGTAGATGAGAGAGATCGATTCGGAAGAAATTTGAGAAAGTAGAT 191043
Qy 7424 AGACTTCTGATATTTATGAAAAAAATGTAGATTTTCAACAGAGGCAAGAAATATATCA 7483
Db 191044 CAGCTATATGAAAAAATAGTAGAGAAAGAACATCTGTGTGAAGAGTTTATGAGAAAG 191103
Qy 7484 CTTTACTGTGTGGTGGGAATCGAGGCTTGGCTTCGCGATCAGGAACAGTGGCAGTTA 7543
Db 191104 AAGGATTTGTTATTTGAAGAAATATAGAAAGAGGTCGTTATTTGAACAAGTTGAAGAT 191163
Qy 7544 CAAATATTTAAAGAAATTCGAGAGTTACTGTTGAAAAATCTTTTGTGAAGCAGCTGAAA 7603
Db 191164 CA-----AAACAGTTCTGAAAAAATCAGAAATATCATCAGTATTTGAAGAGTTAAAG 191217
Qy 7604 AAGTAATGTTAGATCGGATATTTACAGAAATGTTGCTTTACAGCAATCAAGTCTCG 7663
Db 191218 AAGTAAAGAAAGAAATCAGTATAGTAAAAAATAGTAGAAAAAGAAATCTGTTAATCTG 191277
Qy 7664 TAGGAGATTTGGGAATAGAGCTGCTATGCAAAATTAATTTCTAATGGAAGATCAAAAT 7723
Db 191278 AAAAAATTTGACACAAAGAAATCTAATCTAATAAAAAATAGTAAAGATGTATCA--- 191333
Qy 7724 TCAGTATTTAAATTTCTAAGCTATTTAGAAAAAATATATGATGTTATTTGAAGATTAAT 7783
Db 191334 --CTTACTGAAGATTTTGTGAACAAACAGATTCAGTACTGAAAAAGTTATGAAACAG 191391
Qy 7784 CGGAATTTGAGAGCGGAAACAAAGATTAACCTAGAGCGGTAGCTGCCGAGCCATTA 7843
Db 191392 AAGGATCAAAACACTGAAGTTGACAGAGACGTTGAAAGAAAAAGATCCGCTTCTGATGAC 191451
Qy 7844 TCTCAAAAGCAAGAAATGAATTAATTCAGAGGTGAATTTGAAGAGATATTTTCATG 7903
Db 191452 ATGAACAAAGAAATGTATCATGTAATGACAAAGTTACATATGAGAAAAATTCAGTTACT 191511

Qy 7904 AAGAA--AATAGATTAAGTACCCCTTCTAAAGAAATTTGGAAGAAATCAATGTC----- 7957
Db 191512 AAGAAATGTGACAGAGATGCTCTGCTACTGAAAGAAATGTGAAAGAAATATGATCAGTTA 191571
Qy 7958 -----AAGTGAAGAAAAAGAAACAGAGTACTGCTGAATCTCAAGAGCTTCTGTAGAG 8011
Db 191572 CGGAAGAGTGAACGAACAGATCCGTTACTGAAAGAAATATATGAAAGAACACAGTTA 191631
Qy 8012 CAGTACGAGGGCAGGAATTAATTTCCGAAGCAAAAGATCCGGAAGCTCTTATTTGAAG 8071
Db 191632 CTGAAGAGATCGTGAAGAAATGATGATCACTGAAGAAAGTGTGAAGATGATCAGTTA 191691
Qy 8072 TTATGACAAAAATCCGGAAGAAATTTTTCATGCAATATGTAATATGAAAGCAAC 8131
Db 191692 TTCAAGAAATGTGAAAGATGATGATCACTGAAAGAAATTAAGTACAAAGAAATGATCAG 191751
Qy 8132 ATAAAAATGAAGTAAACAGCAGTTCTTAAGCAATTAACAGTTCGTATTTGGAGAGTTG 8191
Db 191752 TTACTGAAGAAATATAGTAGAAGAAAGATCAGTTAATGAAGATTTGAAGAAAGTAT 191811
Qy 8192 GAGTCAACAGGACAGAGCTACTGCTGCAAGTAAACTATGTTAGAAAGTTGAGAGAA 8251
Db 191812 CAGTTTCTGGAAGTGTATGAACAGAAATATGTTACTGAAGAAATGAGAAAGAGAT 191871
Qy 8252 ATTGTTCAAGAACAAATGCAATTTGCAATTTTC--TAAAGTAGAGGTTTGGATGAAG 8308
Db 191872 CAGTTGTTGAAGAAATTTGTAAGAGAGAAAGATCAGTTCTGGAAGAGTTGATGAACAG 191931
Qy 8309 ATTAAGTAACTGCTAAATCTTCTGTATATCAGAAATGAGAGAAATTTGCGAGCAG 8368
Db 191932 AATATGTTATCTGAAGAAATAGTAAGAAAGAAAGATCAGTTGTTGAAGAAATTTGAGAGAG 191991
Qy 8369 GAGTAATATCTTCAACAGCAAAAGTAACTGAATCCGTACTGTTTACAAAGCAAG 8428
Db 191992 AAGGATCAGTTTCTGGAAGAGTTGATGAACAGAAATATGTTACTGAAGAAATGAGAGAG 192051
Qy 8428 A-----TATGAAATAATGATTCACAAAAAATAATTTCAAGATCAATGCTC 8479
Db 192052 AAGGATCAGTTGTTGAAGAAATTTGTAAGAGAGAGATCAGTTCTGGAAGAGTTGATG 192111
Qy 8480 TTGCTTTAAATGATCAAAAGAAATGGAAGCGAATATGAATCTTTAGCCGTAGCCGTGTC 8539
Db 192112 AAACAGAAATGTTTACTGAAAGATGAAGAAAGAAAGAGATCAGTTGTTGAAGAAATTTGAG 192171
Qy 8540 ATGCAACAGAACAAACAAAGCAATTTACGAGATCAACAGATTAATCTTACACTGTA 8599
Db 192172 AGGAGAAAGATCAGTTGTTGAAGAAATTTGAGAGAGAGATCAGTTCTGAAGTAG 192231
Qy 8600 ATGAGAGAAAGTAT-----CTCACTTCGTGCAAAAGCTTTGGCTAAAAATGAA 8650
Db 192232 TTGATGAACCTGAATTAATGATTAATGATGAAGAAATGTAAGAAACAGCTCAATTTACTGAAGAG 192291
Qy 8651 ATTATGGAATGTTAAAGGAATCGAGAGCCCTTAGTCCGAGCGGAACAGCAGCCGTTG 8710
Db 192292 TTGAAGAACAGATATCAGTTTAACAGATGAATTAATAGAGATGATATGCTGAAGCAG 192351
Qy 8711 AAAATTTATCAAGAGTACTACAGAGCAATGTTGTCAGAGAAATTTGGAAATTTGAGAT 8770
Db 192352 TTGAAGAAATCAGATATTAATGATCAAGTATCCCAAGAA-----GAAGAAACAGAAA 192405
Qy 8771 AATTGAAGACATTTGCAAGAGATATAGATTTGTAAAGTCAACGAGACGGAACCAAG 8830
Db 192406 AAGGATTTGTTATTTGAAAAAGTTGAAGAAACAGAGCCGTTTCTGAAGAAATGTTCAAG 192465
Qy 8831 GAGGCTTTCGGAAGAAATGATTTCTGTGAAAAATCAATTTTCAGGGGAAACAAAT 8890
Db 192466 ATGATTAATTTACAGAGAAATTTTGAAGAAATCTGAATCCGTTAATGGGAAATTAATA 192525
Qy 8891 CATTCATTAAGATTAAGCAGAAATTTGGAACCGGAAGTGAATGTAGATGCTTTGA 8950
Db 192526 ACBAAGATCAGATGCTGAGAAATTTTGAACCAAAATTTCTTAAGAAAGATTTAG 192585

Query Match 1.2%; Score 120.2; DB 14; Length 38662;
Best Local Similarity 41.0%; Pred. No. 6.9e-07;
Matches 1172; Conservative 0; Mismatches 1658; Indels 27; Gaps 3;

QY 6109 GACCTTACAAAAAGAAAAATAGTAATPAAAAAGATTTGTTACCAATAGTTCAGCT 6168
DB 31019 GAGTTAATACAAAGATACGAGAGATAGTAATACAAATAGTTACTGAGCAGTGCAGTA 30960
QY 6169 ACTCATACTTTTAAATCTTTATTTGGCAAAATGCCGCTGTTTACGACAAAGCCGAGTGC 6228
DB 30959 GCAGCGCGAGTAAATAGTAGTAGTAGAGCAGGTTGCAAGTAGCTGTGGCAGAGAGTA 30900
QY 6229 GGAATCTTTAATATCAACAAGTTTATGGAAGAAAGAAAGCTCTTGTAGAAAAATCTATA 6288
DB 30899 GCAGCAGCAGCAGTTCCTGAGCTGTGGCAGAGGAGTAGCAGCAGCAGTTCCTGTA 30840
QY 6289 TTAATATGCAAAACATTATTTCTGTAAATAGAGGATTTACAGAAATTCATTCGAGTAGTA 6348
DB 30839 GCTGTGGCAGAGAGATAGCAGCAGCAGGTTGCCGTAGCTGTGGCAGAGAGAGTAGCA 30780
QY 6349 GGTTCGTGTGTGTGTGGTGAATATAGAGATAGAGCTTCTTGTATACCAATATTATA 6408
DB 30779 GCAGCAGCAGTTCCTGAGCTGTGGCAGAGAGTAGAGCAGCATCAG-----CAGTTGCA 30726
QY 6409 AAAAGAAATACCAAGACAAAGAGTTGAAAACTACAATGTCTGTATGAAGTTTCGAGAA 6468
DB 30725 GTAGCTGTGGCAGAGTAGCAGCAGCAGCAGTTCGAGTAGCTGTGGCAGAGAGAGTAGCA 30666
QY 6469 GAAAGCTGAATTAACAGCAGATTTCTAAGCAAGAAATTTCTTTTGGAGTCGAGATTCGA 6528
DB 30665 GCATCAGCAGTTCAGTAGTCTGTGGCAGAGAGTAGAGCAGCATCAGCACTTCAGTAGT 30606
QY 6529 GCAGCCGAGGTAGAGCCGAGGTGGCAGGAACCGTTCCGTAAATCAATTTGCAAGAAAG 6588
DB 30605 GTGGCAGAGAGTAGTAGCAGCATCAGCAGTTCGAGTAGCTGTGGCAGAGATAGCAGCAGCA 30546
QY 6589 ACGGAATAGATGTGGAAGAACAAAGATTTTGTAAAAAAGCTGAGATTACAGCAAAA 6648
DB 30545 GCAAGTTGCAAGTAGCTGTGGCAGAGAGTAGCAGCAGCAGTTCGAGTAGCTGTGGCAGCA 30486
QY 6649 CGTTATAGTTCGTGTGCAATTTGGAAAATGCCGAGTCCGAGTGGCTGCCAAAAGAGACTGGA 6708
DB 30485 GTAGCAGCAGCAGCAGTTCAGTAGTGTGGCAGAGAGTAGAGCAGCATCAGCAGTTGCA 30426
QY 6709 ATTGAGCAGCAGTGGCAGTTTACCAAGATGATCAACACAGAGCAAGATGAAAAAT 6768
DB 30425 GTAGCTGTGGCAGAGAGTAGAGAGAGTTGCAAGTAGCTGTGGCAGAGATAGCAGCAGCA 30366
QY 6769 TCTTAAATATATGACTCGAAAACAAGTTAGATGTATATGCAAAAAATGAGTTAAATCAGGT 6828
DB 30365 GCTGTGTGCAAGTAGCTGTGGCAGAGAGTAGAGCAGCAGCAGTTCGCTAGCTGTGGCA 30306
QY 6829 ACTGGAATGGGTTTACGCCGAGCTGGAAATTTCTTGACCCGAGATCTGGAGATGTTCT 6888
DB 30305 GGAAGATTAGCAGCATCAGCAGTTGCAAGCTGTGGCAGAGTAGAGCAGCATCAGTAGTT 30246
QY 6889 GTCAATTAATATTGCAAAATAGGTAGAAACAGATATGATCATAGTACTTTACCTTCT 6948
DB 30245 GCAGTAGCTGTGGCAGAGAGTAGAGAGAGCAGCAGGTTGCCGTAGCTGTGGCAGAGAGA 30186
QY 6949 ACTGATGTATAATGTAAAGCTCTTAATAAAATTTGCAATTCCTTACAGCCGAGTAGGA 7008
DB 30185 GTAGCAGCAGCTTGCAGTAGTCTGTGGCAGAGAGTAGAGCAATAGCAGATTGCAAGTAGCT 30126
QY 7009 GCCGAGGCTTTCAGCAGCAGTTTACCGAGTGTGTTTCTGTTAACATATTAATAGTTCTGTG 7068
DB 30125 GTGGCAGAGAGTAGAGCAGCATCAGCAGTTGCAAGTAG-----CTGTG 30084
QY 7069 ATAGCTCGAGTTTCACAATTAATCTGTGATTTGACTTCCGTACGAGAAAAATTAATGTAAAG 7128
DB 30083 GGAAGAGAGTAGTAGCAGCATCAGCAGTTGCAAGTAGCTGTGGCAGAGAGTAGCAGCAGTCA 30024
QY 7129 GCAAAAAGAGAAAAAATATTAAGCAAAACAGACCAAAATGACGAATCGAGAGACAGCA 7188

DB 30023 GCAGTTGCAGTAGCTGTAGCAGAGTAGCAGATCAGTAGTTGCAAGTAGTGTGCAGAGA 29964
QY 7189 ATCCGAGCCAAATGCTTGTGTAAATTAATTTTGGAAACAGCTGTAGAGATGAAAAATCT 7248
DB 29963 GAGTAGCAGCAGTTCAGTAGTCTGTGGCAGAGAGAGTAGAGCAGGTTGCAAGTAGCTGTG 29904
QY 7249 GAAGAAAAAGCAAGAGTTTAAATACTTTAGACGAAGTTTAAACAAAGAAACAGATTA 7308
DB 29903 GCAGAGAGAGTAGTAGCAGCAGTTCAGTAGTGTGGCAGAGAGAGTAGAGCAGCATCAGCAGTT 29844
QY 7309 AAGATTAATGATGCTACGAAAAAAATCTTACATATGCAAGTATTTCTACAGAGATACT 7368
DB 29843 GCAGTAGCTGTGGCAGAGAGTAGAGCAGCATCAGCAGTTCAGTAGTGTGGCAGAGATA 29784
QY 7369 TCTGTAAAAAGGAGTAGAGAGATCTCAGAGAGAGAGAAATTAAGCCATGTTGAAGCT 7428
DB 29783 GCAGCATCAGCAGTTCAGCAGCTATGACAGAGAGAGTAGAGCAGCATCCGAGTTGCAGCA 29724
QY 7429 TCTGATTAATTTGAAAAAATGTAGATTTTCAACAGAGAGCAAGAAATTAATCACTTCT 7488
DB 29723 GCTGTGGCAGAGAGTAGTAGCAGCATCAGCAGTTCAGCAGCTGTGTGGCAGAGATAGCAGCA 29664
QY 7489 ACTGTGCTTGGGAACTGCAGAGTCTTCCGATCAGAAACAGTGGCAGTTACAAAT 7548
DB 29663 TCAGCAGTTGCAAGTAGTCTGTGGCAGAGAGAGTAGAGCAGCAGGTTGCAAGTAGCTGTG 29604
QY 7549 ATTAAAGAAATTCGCGAGTTTACTGTGAAAAATTTTGTGAAAAGCAGCTGAAAAAGTA 7608
DB 29603 GCAGAGAGAGTAGAGCAGCAGTAGTAGCAGTTCAGTAGCTGTGGCAGAGAGTAGAGCAGCATCA 29544
QY 7609 AATGTAGATGAGATATTATACAGAAATTTGCTTTAACAGCATATCAAGTCTCTTAGGA 7668
DB 29543 GCAAGTTGCAAGTAGCTGTGGCAGAGAGTAGTAGCAGATCAGCAGTTGCAAGTAGCTGTGCA 29484
QY 7669 GCATTGGAAATGAGAGCTGCTTATGCAAGATTTAAATCTTAATGGAAGATCAAATATCAGT 7728
DB 29483 GAGAGAGTAGAGCAGCATCAGTAGTTGCAAGTAGCTGTGGCAGAGAGAGTAGAGCAGCAGCA 29424
QY 7729 ATTAAAAATTTCTAAGCTATTGGAAGAAAAATTTGATGTTATTTGTAAGATTAATCCGAA 7788
DB 29423 GTTGCAAGTAGCTGTGGCAGAGAGTAGTAGCAGATCAGCAGTTGCAAGTAGCTGTGGCAGGA 29364
QY 7789 TTGAGAGCGGAAGCAAAAGAGTTAAACCGTAGAGCGAGTAGTCCGAGCCATTAATCTCA 7848
DB 29363 GTAGCAGCAGCAGCAGTTCAGTAGTAGTGTGGCAGAGAGAGTAGAGCAGGTTGCAAGTAGCT 29304
QY 7849 AAAGCAAGAAATGAATGAATTCAGAGTTGAAAATTGAGAAAGTATTTTCAATGAAGAA 7908
DB 29303 GTGGCAGAGAGAGTAGAGCAGCAGTTCAGTAGCTGTGGCAGAGAGAGTAGAGCAGCAGCA 29244
QY 7909 AATAGAGTAATAGCCCTTTCAAGAGATTTGGAAGAGAAATCAATGTCAAAGTGGAAAAA 7968
DB 29243 GTTGCAAGTAGCTGTGGCAGAGAGTAGTAGCAGCAGCAGGTTGCAAGTAGCTGTGGCAGGA 29184
QY 7969 GAAAAACAGAGTAGCTGCTGAATTTCAAGAGACTTCTGTAGAGAGCAGTAGAGCAGAGGACAGA 8028
DB 29183 GTAGCAGATCAGCAGTTCAGCAGCAGCTGTGGCAGAGAGAGTAGAGCAGCATCAGCAGTTGCA 29124
QY 8029 ATTATTTCCGAAGCAAAAGATGCCGGAAGCTTTATTTGAAAAGTTAGTACAAAATCCGGA 8088
DB 29123 GCACCTGTGGCAGAGAGTAGTAGCAGCATCAGCAGTTGCAAGAGCTGTGGCAGAGAGAGTA 29064
QY 8089 AGAAGTATTTTTCATGCGAGATTAAGTGAATATGGAAGCAACATTAATAATGAAGTAA 8148
DB 29063 GCAGCAGCAGAGGTTGCGAGCAGCTGTGGCAGAGAGAGTAGAGCATTCGAGGTTGCAGTA 29004
QY 8149 GCAGTTTCTAAAGCAGTAACAGGTTCTGTATTTGGAGAGAGTTGAGAGTCAACCAAGCAGAA 8208
DB 29003 GCTGTGGAGAGAGTAGTAGCAGCATCAGCAGTTGCAAGTAGTGTGTGGCAGAGAGTAGAGCA 28944
QY 8209 GCTACTGTGCAAGTAAACATATGTAGTAAGATTGAGAAAGAAATTTGTTCAAGAACAAAT 8268

Dh	28943	GCAGCAGTTGGCAG-...TAGCTGTGTGATAGAGATGACAGCATCAGTATGTTGCAGTAGCTGTG	28887
Qy	8269	CGATTGAATCAATTTCTTAAAGTAAAGTTTGGATGAAAGTAAAGTAACTGCTAAATTT	8328
Dh	28886	GCAGAGAGAGTACAGCATCAGCATGAGTGGCAGTACGTGTGTGGCAGAGAGATACAGCATCA	28827
Qy	8329	TCGTGTAGTATCAGAGAAATGAGAGAGAAATTGGCCGGAGCAGGAGTGAATACTTCTACAGA	8388
Dh	28826	GCAGTTGCATTACTCTGTGGCAGAGAGAGATGACAGCATCAGCAGTTGCAGTACGTGTGGCA	28767
Qy	8389	CAAAAGTAATACTGAATCCGTAGTCTCTGTTTTCGAAACCAAGATTATGAAATATGATTAC	8448
Dh	28766	GGAGAGATACAGCATCAGCATGACAGTTGCAGTACGTGTGGCAGAGAGATACAGCATCAGCA	28707
Qy	8449	ACAAAAAATATATTTTCAGAAAGTCAATGCTCTTGCTTTTAAATGATATCAAAAGATGAAGCG	8508
Dh	28706	GTTGCAGTAACTCTGTGGCAGAGAGATGACAGCATCAGCAGTTGCAGTACGTGTGGCAGAGGA	28647
Qy	8509	AATATTAATCTTTAGCGGTAGCCGGTGTGCATGCAACAAGAACAAAGCATTTTAG	8568
Dh	28646	GTAGCAGACACAGCATGTTGCAAGTACGTGTGTGGCAGAGAGATACAGCATCAGCAGTTGCA	28587
Qy	8569	AGATCAAAACAAGTTAATCTTCAACACTGTAAATGAGAGAAACGTATCTCAACTTCGTGCA	8628
Dh	28586	GTACCTGTGGCAGAGAGATGACAGCATCAGCATGTTGCAGTACGTGTGGCAGAGAGATGA	28527
Qy	8629	AAAGCTTTGGCTTAAAAATGAAAAATTATGAGAAATGTAAAAAGAACTGAGAGAGCTTATGTC	8688
Dh	28526	GCAGCATCAGCAGTGTGACGTAGTCTGTGGCAGAGAGATGACAGCATCAGTATGTTGACATGCT	28467
Qy	8689	GGAGCGGAAAACAGACGCCGTTGAAAAATTATCAAAAGATATCTACAGAGCATGTTGCA	8748
Dh	28466	GTGCGACAGAGAGATGACAGCAGCAGCTGTGTGTGTGGCAGAGAGAGATGACAGCA	28407
Qy	8749	GGAAATTTGGGAAATTTGAGATTAATTTGAAACGATTTGCAAGAGATATACGATTGTAAAG	8808
Dh	28406	GTTGCAGTAGCTGTGGCAGAGAGATGACAGCATGTTGCATTACTGTGGCAGAGAGATGA	28344
Qy	8809	GTCAACGAGACGGAACCAAGAGAGTCTTGTGGAAAGAAATGTATTTCTGTGAAAAAT	8868
Dh	28346	GCACATCAGCAGTGTGCAGTACGTGTGTGGCAGAGAGATACACAGCATCAGCAGTTGCAGTA	28287
Qy	8869	ACAAATTTGAGGGGAAACAAATATCATCTTGAAGATTAAGCCAGAAATTTTGAACCGGA	8928
Dh	28286	GCTGTGGCAGAGATGACAGCATCCGCAAGTTTGCACAGCTGTGGCAGAGAGATGACAGCA	28227
Qy	8929	AGTGTAAATGTATGATGCTTTGAATGAAGAACTTGATGATG	8965
Dh	28226	TCGGCAGTTGCAGCAGCTGTGGCAGAGAGATGACAG	28190
RESULT 15			
AC138510/c	AC138510	36348 bp	DNA linear HTG 29-JAN-2003
LOCUS	AC138510		
DEFINITION	Homo sapiens chromosome 16 clone LAL6c-448D5, WORKING DRAFT		
ACCESSION	AC138510		
VERSION	AC138510.2	GI:28008372	
KEYWORDS	HTG; HTGS; PHASE2; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 36348)		
AUTHORS	DOE Joint Genome Institute.		
TITLE	Sequencing of Human Chromosome 16		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 36348)		
AUTHORS	DOE Joint Genome Institute.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-JAN-2003) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		

REFERENCE	3 (bases 1 to 36348)
AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submision
JOURNAL	Submitted (29-JAN-2003) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On Jan 29, 2003 this sequence version replaced gi:27545011.

This entry has been annotated with sequence quality estimates computed by the PIRAP assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Bases by base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Sequence Quality Assessment: This entry has been annotated with sequence quality estimates computed by the PIRAP assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 36348: contig of 36348 bp in length.

```
location/Qualifiers
1..36348
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="LA16C-44BD5"
```

ORIGIN			
Query Match	1.2%	Score 119.8	DB 14
Best Local Similarity	41.6%	Pred. No. 8e-07	Length 36348
Matches 942	Conservative	0	Mismatches 1297
			Indels 24
			Gaps 2

QY	6109	GACTATTCACAAAACAAAATAAAGTAATATAAAAAAGATTGTTCCATATGTTACAGT	6168
Db	17227	GAGTTAATGACAAAGATACGGAGAAATGTAATATAAATATGTTACATGACGTGCAGTA	17168
QY	6169	ACTCAATCTTAAATCTTTATTATGGCAATGCCGCTGTTTCAGGACAGCCCGAGTGGCA	6228
Db	17167	GCACGGCAGTAATATAGTATGTAATGTAAGACAGCGATTGCAGTAGCTGTGGCAGAGGAGTA	17108
QY	6229	GGAATCTGTAATATCAACAAGTTTATGAGAAACAGAAGCTCTTGTAAGAAATTTCTATA	6288
Db	17107	GCACGACAGCAGATTTGCCATGCTGTAGCTGTGGCAGAGAGTATGACAGACACAGTTGGCCGTA	17048
QY	6289	TTAAATGCAAAACATTATTTCTGTAAATCAGAGATTAACAAGCAATTCATCCGAGTATGTA	6348
Db	17047	GCTGTGGCAGAGAGAGTAGACAGCAGCGAGTTGGCCGTAGCTGTGGCAAGAGAGTAGCA	16988
QY	6349	GATTCTGTGGTGTGTGTGCAAAATGTAAGAGTAGAGCTTTCTTGATACCAATATTATA	6408
Db	16987	GCAGCAGCAGTTGCCATGCTGTAGCTGTGGCAGAGGAGTAGACAGCATTCAGC-----AGTTGCA	16934
QY	6409	AAAAGCAATATCAAGACAAAGATTGGAAAAACATCAATGTCTGATATGAAGTTTCCGAGAA	6468
Db	16933	GTAGCTGTGGCAGGAGTAGCAGCAGCAGCGAGTTGCAGTACGCTGTGGCAGAGAGATGACA	16874

QY 6469 GAACTGAAATTAACAGCAATCTTAAGCAAGAAATTCCTTTTGGAGTCGGATCGCA 6528
 Db 16873 GCATCAGCAGTTGCGACTAGCTGTGGCAGAGAGAGTACAGCATTCACAGATTGGAGTACGT 16814
 QY 6529 GCAGCCGGGGTGGAGCCGAGTGGCAGGAACCGTTCCGTAAATCAATTTTGCAGAAAG 6588
 Db 16813 GTGGCAGAGAGATTACAGCATCAGCAGTTGTCAGTACGTGTGGCAGAGATTAGCAGACGA 16754
 QY 6589 ACGAGATGATGTGGAAGAAAGCAAAAGATTTGGTAAAAAAAGCTGAGATTACAGAAA 6648
 Db 16753 GCAGTTGCAGTACGTGTGGCAGAGAGTACAGCAGCAGCAGTTGTCAGTACGTGTGGCAGGA 16694
 QY 6649 CGTTATAGTCTGTGTGCAATTGGAAATGCCGACGTGGAGTGGCTGCAAAAGAGCTGGA 6708
 Db 16693 GTAGCAGCAGCAGCACTTGCAGTACGTGTGGCAGAGAGTACAGCATTCAGCAGTTGCA 16634
 QY 6709 ATTGGAGCAGCAGTGGCAGTTACCAAGATGAAATCAAAACAGAGCAAGATGAAAAAT 6768
 Db 16633 GTAGCTGTGGCAGAGAGATTAGCAGCAGTTGTCAGTGTGGCAGAGATTAGCAGACGA 16574
 QY 6769 TCTAAATTTATGACTCGAAACAAGTTAGATGTATATGACAGAAAATGAGATTAATCAGGT 6828
 Db 16573 GCTGTTGCAGTACGTGTGGCAGAGAGAGTACAGCAGCAGCAGTTGCCGTAGCTGTGGCA 16514
 QY 6829 ACTGGAATGGTTCAAGCCGAGCTGGAATTTCTTGACCCGAGATATCTGAGTGTCT 6888
 Db 16513 GGAGATTAGCAGCATCAGCAGTTGCAATGCTGTGGCAGAGTACAGCATCAGTAGTT 16454
 QY 6889 GTCAATTAATTTGCAAAATAAGTAGAAGATATGATCATAGTCTTTACACTCTTCT 6948
 Db 16453 GCAGTAGCTGTGGCAGAGAGATTAGCAGCAGCAGCAGTTGCCGTAGCTGTGGCAGAGGA 16394
 QY 6949 ACTGATGTAATGTAATAAGCTCTTAATAAAATTCGATTCCTTGACCCGCTGGAGGA 7008
 Db 16393 GTAGCAGCAGTTGCGACTAGCTGT-----GGCAGAGAGATTAGCAGCA 16352
 QY 7009 GCCGACGCTCTTGACAGCAGTTACCGAGTGTCTTCTTAACCATTAATAGTTCTGTG 7068
 Db 16351 ATAGCAGTTGACAGTACGTGTGGCAGAGAGATTAGCAGCATTCAGCAGTTGCAATAGCTGTG 16292
 QY 7069 ATAGCTCGAGTTCACAATACTCTGATTGACTTCCTGACGAGAAAATTAATGTAAAG 7128
 Db 16291 GCAGAGAGAGTACAGCATCAGCAGTTGCAATGCTGTGGCAGAGAGATTACAGCATCA 16232
 QY 7129 GCAAAAGAGAAAAAATTTAAAGCAACAGCAGCAATGCAAGATCGAGAGAGAGA 7188
 Db 16231 GCAGTTGACAGTACGTGTGGCAGAGAGTACAGCATCAGTGTGCACTAGCTGTGGCAGGA 16172
 QY 7189 ATCGAGCCCAATGTCTTGTTAAATAATTTTGAACAGCTGTGAGAGATGAAAAATCT 7248
 Db 16171 GAGTAGCAGCAGTTGCAATGCTGTGGCAGAGAGAGTACAGCAGTTGCAATAGCTGTG 16112
 QY 7249 GAAAGAAAAAGAACAGAGTTTAAAACTTTAGAGAGATTAAACAAAGATTA 7308
 Db 16111 GCAGAGAGAGTACAGCATCAGTACGTGTGGCAGAGAGAGTACAGCATTCAGCAGTT 16052
 QY 7309 AAGTAAATGATGCTACGAAAAAATCTTAACATCAGCAGTATTTCTACAGAGTACT 7368
 Db 16051 GCAGTAGCTGTGGCAGAGAGAGTACAGCATCAGCAGTTGCACTGTGGCAGAGAGTA 15992
 QY 7369 TCTGTAAAACCGGATAGAGAGATCTACAGGAGAAAGAAATTAAGCCATTTGTGAAGCT 7428
 Db 15991 GCAGCAGTCGGCAGTTGCACTAGTACGAGAGAGAGTACAGCATTCGACAGTTGAGCA 15932
 QY 7429 TCTGATATTATTGAAAAAATGTAGATATTACACAGAGAGACAAAGATTAATCACTTCT 7488
 Db 15931 GCTGTGGCAGAGAGATTACAGCATCGCAGTTGTCAGCAGCTGTGGCAGAGATTAGCAGGA 15872
 QY 7489 ACTGTGTTTGGAGACTGCAAGTCTTCTCCGATCAGGAAACAGTGGCAGTTCACAAT 7548
 Db 15871 TCGGCACTTGGCAGTACGTGTGGCAGAGAGAGTACAGCAGCAGCTTCACTAGCTGTG 15812
 QY 7549 ATTAAAAAGAAATTCGGAGTTACTGTGAAAAATCTTTTGTGAAAGCAGCTGAAAAAGTA 7608

Db 15811 GCAGAGAGAGTACAGCAGTACAGTTCAGTACGTGTGGCAGAGAGAGTACAGCATCA 15752
 QY 7609 AATGTTAATCGAATATTACAGAAATGTGCTTTAAACAGCATCAAGCTCTCTAGGA 7668
 Db 15751 GCAGTTGCAATGCTGTGGCAGAGAGAGTACAGCATCAGCAGTTGTCAGTACGTGTGGCA 15692
 QY 7669 GCATTGGAAATAGAGCTGCTATGCAAAATTTAAATCTTAATGAAAGATCAAAATATCAGT 7728
 Db 15691 GGAAGAGCAGCAGCATCAGTATGTTGCACTAGTACGTGTGGCAGAGAGAGTACAGCAGCA 15632
 QY 7729 ATTAAAAATTTCTAAGCTATTAGAAAAAATTTGATGTTATTGTAATAAGATTAATCGGA 7788
 Db 15631 GTTGCAGTAGCTGTGACAGAGTACAGCAGCATCAGCAGTTGTCAGTACGTGTAGCAGGA 15572
 QY 7789 TTGAAGCGGAAAGCAAAAGATTAAACGTTAGAGCGGTAGCTGCCGAGGCCATTATCTCA 7848
 Db 15571 GTAGCAGCAGCAGCAGTTGTCAGTACGTGTAGAGAGAGAGTACAGCAGCAGCAGTACGA 15512
 QY 7849 AAAGCAAAAGAAATGAATGAATTCAGAGGTTGAAATTTGAAAGAGATTTTCAATGAGGA 7908
 Db 15511 GCAGCTGTAGCAGAGAGTACAGCAGCAGCAGCAGAGTACAGAGAGAGTACAGCAGCA 15452
 QY 7909 AATAGATTAATGAGCCCTTTCTAAAGAAATTGGAAGAGAAATCAATGTCAAAGTGA 7968
 Db 15451 GCAGTTGCAATGAGCTGTGGCAGAGAGATGACAGCAGCAGCAGTTGTCAGTACGTGTGCA 15392
 QY 7969 GAAACAGAGTACGCTGGAATCTCAAGAGCTTGTAGAGAGCAGTACAGGCGCAGGA 8028
 Db 15391 GAGAGAGTACAGCATCAGCAGTTGCAATGCTGTGGCAGAGAGAGTACAGCATCA 15332
 QY 8029 ATTATTTCCGAAGCAAAAGATGCCCCGAAGCTCTTAATTTGAAAGTTAGTACAAATCCGA 8088
 Db 15331 GTTGCAGTAGCTGTGGCAGAGAGAGTACAGCATCAGTATGAGTTGCACTAGCTGTGGCAGGA 15272
 QY 8089 AGAAGTATTTTTCATGCAATATGTGAATGTGAAGCAACATTAATAAGTAAACA 8148
 Db 15271 AGAGTAGCAGCAGCAGCAGTTGCAATGCTGTGGCAGAGAGAGTACAGCATCA 15212
 QY 8149 GCAGTTTAAAGCAGTAAACAGTTCTGTTATTTGGAGAGAGTTGGAGTCAACCAAGCAGGA 8208
 Db 15211 GCAGTAGCTGTGGCAGAGAGTACAGCAGCAGCAGCAGTTGCAATGAGTGTGGCAGAGAGTA 15152
 QY 8209 GCTACTGCTGCAAGGTAATACTATGTAGAGTTGAGAGAAATTTGTTCAAGAACAAAT 8268
 Db 15151 GCAGCAGTTGCAATGAGTACGTGTGGCAGAGAGAGTACAGCATGTTGCACTAGCTGTGGCAGGA 15092
 QY 8269 CGATTGAATGCAATTTCTAAAGTAGAGCTTTGATGAATTAAGTAACTGTCTAAATCT 8328
 Db 15091 GCAGTAGCAGCAGCAGCAGCAGTTGCAATGCTGTGGCAGAGAGAGTACAGCAGCAGATT 15032
 QY 8329 TCTGTAGTATCAGGAAATGAGAGAGATTGCCGAGAGAGAG 8371
 Db 15031 GCAGTAGCTGTGGCAGAGTACAGCATCGCAGTTGTCAGAG 14989

RESULT 16
 PFA929351/c 347050 bp DNA 1linear INV 16-Apr-2005
 LOCUS PFA929351 Plasmodium falciplarium strain 3D7, chromosome 5, segment 1/4.
 DEFINITION Plasmodium falciplarium strain 3D7, chromosome 5, segment 1/4.
 ACCESSION AL929351 AL844504
 VERSION AL929351.1 GI:23504490
 KEYWORDS
 SOURCE
 ORGANISM
 Plasmodium falciplarium 3D7
 Plasmodium falciplarium 3D7
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE
 1 Hall, N., Pain, A., Bartman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooke, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N.,


```
/codon_start=1
/product="RIFIN"
/protein_id="CAD51371.1"
/db_xref="GI:23504493"
/db_xref="InterPro:IPR002858"
/db_xref="InterPro:IPR006373"
/db_xref="UniProt/TrEMBL:O81493"
/translation="MAININHLCSKILFSLQNLITSSVYNNKKNPYTPHTSAIT
SVLNBCDIYMPNPDADDMKSVENKFDRSSRFESEYERMIYTRQKKEQCEQDIQ
KIILDKIEKLEKEVKCCILCGGIGGVAPFGLFGLVNMKKKAAVVAATDVG
IKBAIKVGSIFLADGSTWMTKMINAGNSOKRSIVITLILNKCEDEBALGPL
FKASBEAATSEGEVEPFSGNISRMANANBAAGNANANAMTAVGTCISNPVITSA
IVVITAVLILITLILKTRKKKKKKQKQYTKLNNQ"
gene
complement(39443..40488)
/gene="STEVR", degenerate"
/notes="Apparent single frameshift in exon II"
CDS
/pseudo
/codon_start=1
complement(41129..41403)
/gene="RIF", degenerate"
/notes="RIF, degenerate"
/notes="synonym: PFE0035c"
/pseudo
complement(41129..41403)
/gene="RIF", degenerate"
/notes="RIF fragment"
/pseudo
/codon_start=1
complement(43327..47481,47612..47761))
/gene="MBSA"
/notes="synonym: PFE0040c"
complement(43327..47481,47612..47761))
CDS
/gene="MBSA"
/notes="Similar to Plasmodium falciparum mature
parasite-infected erythrocyte surface antigen TR:O06166
(EMBL:AF056936) (1661 aa) fasta scores: E(): 3.6e-144,"
Query Match 1.2%; Score 119.4; DB 2; Length 347050;
Best Local Similarity 42.2%; Pred. No. 3.9e-07;
Matches 803; Conservative 0; Mismatches 1096; Indels 6; Gaps 2;
```

```

Db 45081 AAAGATAAGTGCACAAAACAGAAAAAGTAAAGATTTAAAAAGAAATAAAAAAA 45022
Qy TTTTGAGAGAGCGTGAAGAAAGTAAATGTTAGATCGGATATTACAGAAATGTCCTTTA 7644
Db 45021 GTTAAAAAAGGTAAAAAAGAAATATTAAGATGAAAGAAAGATTAATGTGATGAGA 44962
Qy ACAGCATATCAAGGTCTGTAGAGACATTGGAAATAGAGGTCTCATGAGAAATTAAT 7704
Db 44961 AAAGAAATTAAGAAAGAAAGATTTAAATGAAAAAGATACCGGAAAAGAAAGATTAAGAGTA 44902
Qy TCTAATGAGATCAATATATCAGTATTAAAAATTTAAAGCTATTAGAAAAAATTTGAT 7764
Db 44901 GAACAGAAAAAAGAAAAAGAAAGTTAAAGAAAAAGAAAGATTTAAAGAAAAAGAA 44842
Qy GTTATGTAAAAAGATAATTCGGAAATTGAGAGCGGAGAAAGATTTAACGTAGAGACG 7824
Db 44841 GTTAAAGAAAGAAAGAAAGTAAAGAAAAAGAAAGTAAAGAAAAAGAAAGTAA 44782
Qy GTAGCTGCCGAGCCATTATCTCAAAAGCAAAAGATGAATTCAGAGTTGAAAT 7884
Db 44781 GAAAAAGAAAGATTAAGAAAGAAAGAAAGATTAAGAAAGAAAGATTAAGAAAGATTAAGAT 44722
Qy 7885 GAGAAAGATATTTTCATGAAGAAATAGAGTAACTAGCCCTTTCTAAAGAAATTTGAAAG 7944
Db 44721 AAAGAGATGAGAACAAAGAAAAAGAAAGAAAGTAAAGAAAGTAAAGAAAGATAC 44662
Qy 7945 GAAATCAATGTCAAGTGGAAAAAGAAACAGAGTACGTCTGAATCTCAAGAGCTTCT 8004
Db 44661 GAAAAAGAAAGTAAAGTATAGACAGAGAAATTAATTAAGAAAGTAAAGAAAGAAAGT 44602
Qy 8005 GTAGAGCAGTAGACAGCGGACAGAAATTTTCCAGAGCAAAAGATGCCGAGAGCTTAT 8064
Db 44601 AAAAAAGAGTAAAGAAAGAAATTAATTAAGTAAAGAAAGTAAAGTATGATAGTACAA 44542
Qy 8065 TTGAAAGTTAGTACAAAATCCGAAAGAAAGTATTTTCATGCAATATGTGAATGTGAA 8124
Db 44541 G---AAATTAAGAACGAAAGATGTTAACGAAAAAGTACCGGAAACAAAGATTAAGGTGATA 44485
Qy 8125 GCAACACATTAATGAAGTAAGTAACAGCAAGTTCTTAAGAGTAAACAGGTTCTGTATTGGGA 8184
Db 44484 GAACAGAAAAAGAAAAAGAAAGTAAAGAAAAAGAAAGTAAAGAAAAAGAAAGAA 44425
Qy 8185 GAGAGTGAAGTACCAAGAGCAGAAAGCTACTCTGAGGTAAATCTATGTTAGAAAGTTAG 8244
Db 44424 GTTAAAGAAAGAAAGAAAGTAAAGAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAG 44365
Qy 8245 GAAGAAATTTGTTCAAGAACAAATCGATGGAATGCAATTTCTAAAGTAAAGTTGAT 8304
Db 44364 GAAAAAGATACCGAAAGCAAGATTAATGTATAGTACAAAGAAATTAAGAACAAAGATGT 44305
Qy 8305 GAAGTAAAGTAACTGCTAAATCTTCTGTAGTATACAGAAATTTGAGAGAGAAATTCGCCGA 8364
Db 44304 AAGCAAAAAAGTACCGAAAGCAAGATTAATGTATGAGAAAGAAAGTAAATTAAGAAAG 44245
Qy 8365 G---CAGAGAGAAATATTCTTACAGACAAAGTAAATCTGATCCGTAGTTGTTACGA 8421
Db 44244 GTTAAAGAAAGATTAAGAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAG 44185
Qy 8422 AAGCAAGTATGAAGAAATATGATTAACAAGAAAAATTAATTTCAAGAGTCAATGCTCTT 8481
Db 44184 AATAGAAAAAATTAAGAAAGAAAGATGTGATAGAACAAAGAAATTAAGAGCAAGATTAAC 44125
Qy 8482 GCTTTAAATGATTAACAAGAAATGAAGCGAATTAAGAAATTTTAAAGCGGTAGCCGCTGTGAT 8541
Db 44124 GAAAAAGTACCAAAAGCAAGATTAAGAAAGTAAAGAAAGTCAAAAGTCAATATAGAT 44065
Qy 8542 GCAAGAGAACAAAGCAAGATTTACGAGATCAAAAGATTAATCTTACAACTGTAAAT 8601
Db 44064 GATTGTAAAAAAGAAAGAAAGTACAGAGAAATCTGAGAGAGAGTCTGAAAGAAAGTCT 44005
Qy 8602 GGAGGAAACGTATCTCAACTTCGTGCAAAAGCTTTGCTAAAGTAAAGTAAATTAATGAAAT 8661
```

```

Db      44004 GAAAGAAGATCTGAAAGAAGTCTGAAAGAAGATCTGAAAGAAGATCTGAAAGAAGATCT 43945
Qy      8662 GTAAAAAGAACTGGAGGACCTTAAGTCGAGCGGAAACAGCGCCGTTGAAAAATTATACA 8721
Db      43944 GAAAGAAGTCTGAAAGAAGATCTGAAAGAAGATCTGAAAGAAGTCTGAAAGAAGATCT 43885
Qy      8722 AAGAGTACTACGAGGACGANTGGTTGAGGAAATTGGGAAATTGGAGATTAATTAGAAACG 8781
Db      43884 GAAAGAAGTCTGAAAGAAGTCTGAAAGAAGTCTGAAAGAAGATCTGAAAGAAGATCT 43825
Qy      8782 ATTGCAAGAGATTAATACGATTTGTAAGAGTCAACGAGACGGAACCAAGAGAGTCTTTC 8841
Db      43824 GATGAAGAAAAAATATCATCAGGTTTGGTACATPAGAAAGAAATTGTAAGAAAAAGAAA 43765
Qy      8842 GGAAGAGATGTAATTTCTGTGAAAAATACAAATTTACGGGAAAACAAATCATCATTTGA 8901
Db      43764 TATTAATTAATGAGAAATTAGAAAGATTTATTAAGAGAAAACAGATGAGAAATATTTT 43705
Qy      8902 GATTAAGCCAGAAATGTTGAAACCGGAGAGTAAATGTAGATGCTTTGAAATGAACTGAT 8961
Db      43704 GAAAGATTAATTAATTCATCAACAAAGAACTAATTTACTTTGAATACATTTCCCAAATATG 43645
Qy      8962 GTAGATCTTACAGAAAAAGTGGTGGCTATGTGTGAATTTGTAAT 9006
Db      43644 TTAATGAAGATTTTCAGAGAGATTTTCACATATATTAAATTT 43600

RESULT 17
AF056936 5420 bp DNA linear INV 10-MAY-2001
LOCUS Plasmodium falciparum mature parasite-infected erythrocyte surface
DEFINITION antigen gene, complete cds.
ACCESSION AF056936
VERSION AF056936.1 GI:3044184
KEYWORDS Plasmodium falciparum (malaria parasite P. falciparum)
SOURCE Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ORGANISM Coppel, R.L.
REFERENCE 1 (bases 1 to 5420)
AUTHORS Coppel, R.L.
TITLE Repeat structures in a Plasmodium falciparum protein (MESA) that
binds human erythrocyte protein 4.1
JOURNAL Mol. Biochem. Parasitol. 50 (2), 335-347 (1992)
PUBMED 1741020
REFERENCE 2 (bases 1 to 5420)
AUTHORS Kun, J.F., Waller, R.L. and Coppel, R.L.
TITLE Plasmodium falciparum: structural and functional domains of the
mature-parasite-infected erythrocyte surface antigen
JOURNAL Exp. Parasitol. 91 (3), 258-267 (1999)
PUBMED 10072328
REFERENCE 3 (bases 1 to 5420)
AUTHORS Kun, J.F.J. and Coppel, R.L.
TITLE Direct Submision
JOURNAL Submitted (01-APR-1998) Humanparasitology, Institute for Tropical
Medicine, Wilhelmstrasse 27, Tübingen 72074, Germany
FEATURES
source
1..5420
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/strain="Palo Alto"
/db_xref="taxon:5833"
join<1..150,281..>5116
/product="mature parasite-infected erythrocyte surface
antigen"
join(1..150,281..5116)
/note="MESA; PfEMP2"
/codon_start=1
/product="mature parasite-infected erythrocyte surface
antigen"
/protein_id="AAC13303.1"
/db_xref="GI:3044185"
/translation="MEVIGCNLCYDKKNMMENEGNKVKVYNSSLKVKMKFCLCTI
ICVFLDIYNCSPYISYSINNNDRIYRLISLTPSPHSLSEINRTDEHLSIR

```

ORIGIN

```

Query Match      1.2%; Score 118.8; DB 2; Length 5420;
Best Local Similarity 41.9%; Pred. No. 2..2e-06;
Matches 1240; Conservative 0; Mismatches 1672; Indels 48; Gaps 7;

```

```

NYIBCLNAPYIDDPMLGWSVVTDRKNNCLQHIKLENOESRRRKOEBENAKDIEEI
RKEKEVYLMKLEEMDESDEVEKAFREILOPILKLRDRTRRKRNVNNGESKEDSKEDT
ESKESGESKETSGETGESKETSGETGESKETSGETGESKETSGETGESKETSGETGESK
TGESKETSGETGESKETSGETGESKETSGETGESKETSGETGESKETSGETGESKETS
KITSPFRETEVNTKITEESKDEKGNKVSOPYNSESNSTSEETKYLAKRENBGRK
LGENVNDASNSNSDPKULTEOENKVSSESEEDNDPREENEKADNRKSKKRS
FFOMLGCNPLCNKNIETDDEETLVVKDAAKKHFLFLEAETKNDKNDKDLGEGD
KEDVKKNDKOKVYLGGDEDEKVKDOKVYLGDEDEKVKDOKVYLGDEDEKVKDOKVY
EKTOKEIKETKVKKKCKKKVKVKGIGKENTBEGDKVGPDIIEEYBEIKKVED
GKENDTEGANDVKPEPIISEVEBEIKGHEVBEIKENTBEGDKVGPDIIEEYBEI
EIKENVEGINKENTBESDKVIGOEIIEEVEBEIKENTBEGDKVGPDIIEEYBEI
EIKOEENKENTLEIKDVIIGOEVIIEEVEBEIKVKKVKGIGKENTBESDKVIGOE
IIEEVEBEIKOEVEBEIKENTBESDKVIGOEVIIEEVEBEIKVKKVKGIGKENT
VKKEVKEVKKVKKVKKVKKVKKVKEIKENKEDVNDKNDKVIIEEVEBEIKVKE
EKVEYKEVEYKEVEYKEVEYKEVEYKEVEYKEVEYKEVEYKEVEYKEVEYKEVEY
IEOEKEVEYKVEYKOTENDKVIIGOEIIEEIKVKKVKKVKKVKKVKKVKKVKKV
EIMEMLKDNDKNDKVIIEEKEVEYKEVEYKEVEYKEVEYKEVEYKEVEYKEVEY
KDKETEOEKEVEYKEVEYKOTENDKVIIGOEIIEEIKVKKVKKVKKVKKVKKV
VIVOEINMEDVNEKOTANKDVIIEEKEVEYKEVEYKEVEYKEVEYKEVEYKEVEY
DKETEOEKEVEYKEVEYKOTENDKVIIGOEIIEEIKVKKVKKVKKVKKVKKV
IVOEINMEDVNEKOTANKDVIIEEKEVEYKEVEYKOTENDKVIIGOEVIIEEVE
VKRVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKV
KEVNRVRRNRKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKVKKV
QSESESESESESESESESESESESESESESESESESESESESESESESESESESE
ESDEKNTSGLVHRNCKKEKKYNNGLSEYKEKQNEYPDEEYIIOGKVNNTNTF
PMAINEDFRREFNHIILSIHEDRDLMBELKILYVLFENPNNMKKRAELDKFSEM
NVNHIOLYVERIRMYEENNAARGRLNTYILDPITFPNVIFPGDTPKPFIDE"

```

```

Qy      6204 TCGTTCAGACAGCCGAGTGGCAGGAACTTTAATATCAACAAGCTTATGAGAAAC 6263
Db      1381 TCGTCAAGAAAGAAATTTCTGAAAGAAACAGAGATGTAACCGGAAAGAAATGAGAAAA 1440
Qy      6264 AGAAGCTCTTGTGAGAAAATTTCTAATTTAAATGCCAAAACATTTTCGTAAAAATCGAGAG 6323
Db      1441 GGCAGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1500
Qy      6324 TTACAGAAATTCATCGAGAGTAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 6383
Db      1501 TAAATTTCCATGTATATATATATATATATATATATATATATATATATATATATATATAT 1560
Qy      6384 AGCTTCTTCTGATACCA--ATAATATATATATATATATATATATATATATATATATATAT 6440
Db      1561 AGATGATGCTATATATATATATATATATATATATATATATATATATATATATATATATAT 1620
Qy      6441 TACAATGTCTGTATGAAGGTTTGGAGAGAGAACTGAAATTTACAGCAGATTTTAAAGCAAG 6500
Db      1621 TGAATAAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1680
Qy      6501 AATTTCTCTTTTGGAGTGGAGTGGCAGCAGCCGGGTTAGAGCCGAGTGGCAGAGAAC 6560
Db      1681 TGAACGAAAGATTAAGATTTTGGAGAGAGAGATTAAGAGATTTTAAAGAAAAAGATGA 1740
Qy      6561 CGTTCCGTAATCAATTTTTCAGAGAAAGACGAAAGTATGTGTGAAGAAAGCAAAAGTTT 6620
Db      1741 TGAACGAAAGATTAAGATTTTGGAGAGAGAGATTAAGAGATTTTAAAGAAAAAGATGA 1800
Qy      6621 GGTAAAAAAGCTGAGATTAACGCAAAAACGTTATGTTCTGTGCAATTTGGAATGCGCG 6680
Db      1801 TGAAGAAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1860
Qy      6681 AGTCGAGTGGCTGCATAAGAGAGCTGGAATTGGAGCAGCAGTGGCAGTTTACCAAGATGA 6740
Db      1861 AGAAAAAAGAGTTTAAAAAAGAGTTTAAAAAAGAGTTTAAAAAAGAGTTTAAAAAAGAGTT 1920
Qy      6741 ATCAAAACAGAGAGCAGAGATTAATTTCTAATATTAATTAATTAATTAATTAATTAATTA 6800
Db      1921 TACTGAAGGTAACGATTAAGATTAAGAGCAGAGATTAATTAATTT-----GAAGAGT 1971

```

Qy 6801 AATAGCAGAAAAATGAGATAAATCAAGTACTGGAATCGGTTCCGCGAGCTGGAATCT 6860
Db 1972 AAAAGAGAAATTTAAAAACAAGTAGAAGATTAAGAAATGATTAAGAGGTAA 2031
Qy 6861 TGCACCGGAGTATCTGGAGTGGTTCTCTCAATTAATTGCAATTAAGGTGAAGAACGA 6920
Db 2032 CGATTAAGGAAAGGCGCCGGAATATATCTGAAGAGGTAAGAAAGAAATTTAAAAACA 2091
Qy 6921 TATCGATCATAGTACTTTTCACTCTTCTCTATGATTAATGTAAAGCTCTTAATAAAT 6980
Db 2092 TGTAGAAGAGTATTTAAAGAAATTAATCTGAAGTAAAGTAAAGTGAAGAGGCGCAGA 2151
Qy 6981 TTGCAATTCCTTGACACCGCGGTGAAGAGCCGAGGCTTTCGACGAGTTACCGAGTGT 7040
Db 2152 AATATTAATCTGAAGAGTAAAGAAAGAAATTTAAAAACATGTAGAAAGAGAAATTAAGA 2211
Qy 7041 TTCTGTTAACATTAATTAAGTCTGTGTATAGCTGAGTTTCACATTACTTGATTTGAC 7100
Db 2212 AAATGATCTGAAAGTAAAGATTAAGTGTAGACAAAGAAATTAATTAAGTGAAGGTAA 2271
Qy 7101 TTCCTAGCAGAAAAAGTAAATGTAAACGCAAAAGGAAAAATTAATTAAACAAACAGC 7160
Db 2272 AGAAGGAATTAAGAAAAATGATATGAAAAATTAAGATTAAGATTAAGACCGAAATGAT 2331
Qy 7161 AGCAATGCGAAGATCGAGAGCAGCAATCGAGCCAAATGCTTGTGAATTAATTTGG 7220
Db 2332 AACTGAAGAGTAAAGAAAAAGAAATTTGAAAAACAAGAAAGAAAGAAATTAAGAAATAT 2391
Qy 7221 AACACTGTAGAAGATGAAGAAAAATTTCTGAAGAAAAAGAACAAAGTTTAAAACTTT 7280
Db 2392 TCTTAAATTAAGATTAATGTAATGTAGACAAAGATTAATTAAGAAAGTAAAAAAGT 2451
Qy 7281 AGACGAAGTTAAACAAGAAACAAGATTAAGAAATTAATGATGCTAGCAAAAAATCTTACA 7340
Db 2452 AATTAAAAAAGATGAAGAAAAAGAAATTAAGAAATTAATCTACT---GAAAGTAAAGATA 2508
Qy 7341 ATCAGCAGGTATTTCTACGAAGATATCTTGTAAAAAGCGATAGAAGAGATCTCAAGG 7400
Db 2509 AGTGAATAGGACAAGAAATTAATGTTGAAGAGTAAAAAGAAATTTGAAAAACAAGTGA 2568
Qy 7401 AGAAGGAATTAAGCATTTGTGAAGACTTCTGATATTAATTGAAAAAATGTGATTAATAC 7460
Db 2569 AGAAGGAATTAAGAAAAATGATTAAGAAATTAAGTAAAGTAAAGTAAAGCAAGAGTAT 2628
Qy 7461 AACAGAGCAAGATTAATATATCACTTCTAGTGTGGTTGGGAACTGCAGGCTTGTCTTC 7520
Db 2629 AAAAGAGATTTAATGAAGAAAGTCCGAAAAACAAGTAAAGTGAACAAAAACAGAAAA 2688
Qy 7521 CGCATCGAGACAGTGGCAGTTACAAATATTAAGAAAAATTCGGAATTAAGTTGAAGA 7580
Db 2689 AGTAAAGAGTAAAGAAAAAGTAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGAAAGA 2748
Qy 7581 TTCTTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGATTCGGAATTAACAGAAATGTTGC 7640
Db 2749 TAATTAAGATTAAGAAAAAGTAAATGTTAAGAAAAAGAAATTAAGAAAAAGAGTTAA 2808
Qy 7641 TTTAAACAGCATATCAAGTCTCTGAGAGCAATTCGGAATAGGAGCTGCTATGCAAGATT 7700
Db 2809 TGAAAAAGATACCGCAAAACAAGATTAAGAGATAGCAAGAAAAAGAAAAAGAAAGAGT 2868
Qy 7701 AAATTTCTAATGAGATCAAAATATCAATTAAGAAATTTCTAAGCTATTAGAAAAAATAT 7760
Db 2869 TAAAGAAAAAGAAAGAA-----GTTAAAGAAAAAGAAAGTAAAGAAAGAAAGA 2916
Qy 7761 TGAATGTTATTTGTAAGAAATTAATCGAATTTGAGAGCGGAAGCAAAAGATTAAACGTAG 7820
Db 2917 AAGAGTAAAGAAAAAGAAAGTAAAGAAAAAGAAAGTAAAGAAAGAAAGAAAGT 2976
Qy 7821 AGCGGTAGCTGCGGAGCATTATCTCAAGCAAAAGAAATGAATGAATTCAGAGTTGA 7880
Db 2977 AAAAGAAAAAGAAAGTAAAGAAAAAGAAAGAAAGTAAAGAAAGAAAGTAAAGAAAGCA 3036
Qy 7881 AATTGAAGAGATTTTCAATGAAGAAAAATGAGTAACTAGCCCTTCTAAAGAAATGG 7940

Db 3037 AGATTAAGAGATAGAACAAAGAAAAAGAAAAAGAAAGTAAAGAAAGTAAAGAAAGAAAGA 3096
Qy 7941 AAGGAATTCATATGTCAGAGTGAAGAAAAAGAAACAAGATGACTGCTGAATCTCAAGAGC 8000
Db 3097 TACCGAAAAACAAGTAAAGTGAATGACAAAGAAATTAATTAAGAAATTAAGAAAGAAAGA 3156
Qy 8001 TTCTGTAGAGCAGTAGAGGCGCAGAAATTAATTTCCGAACCAAAAGATGCGGAGGCTC 8060
Db 3157 AGTTAAAAAAGAGTAAAGAAAAAGAAATTAATTAAGAAATTAAGAAATTAATGATAGT 3216
Qy 8061 TTATTTGAAGTATGATACAAAATCCGAAAGAAATTTTTCATGCAATTAATGTGAATAT 8120
Db 3217 ACAAGAAATTAATGAACGAGATGTTACGAAAGATCCGCAAAACAAGATTAAGTATAGA 3276
Qy 8121 GGAAGCAACACATA-----AAATGAAGTAAACGACATTTCTAAGCAGTAAACAGTTTC 8174
Db 3277 ACAAGAAAAAGAAAAAGAAAGTAAAGAAAAAGAAAGAAAGTAAAGAAAAAGAAAGT 3336
Qy 8175 TGTATTTGGAGAGTTGAGTCAACCAAGCAGAAAGCTACTGCGAGTAAATTAATGAT 8234
Db 3337 AAAAGAAAAAGAAAGTAAAGAAAAAGAAAGAAAGTAAAGAAAAAGTAAACGAAAGCA 3396
Qy 8235 AGAAGTTGAGAAAGAAATTTGTTCAACAATCGATTTGAATTCATTAAGTGA 8294
Db 3397 AGATTAAGAGATGAAACAAGAAAAAGAAAGAAAGTAAAGAAAGTAAAGAAAGAAAGA 3456
Qy 8295 AGGTTTGAATGAATGAATGAATGTAATCTGTAATCTTCTGATAGTACG-----GAAATGG 8348
Db 3457 TACCGAAAAACAAGTAAAGTATGACAAAGAAATTAATTAAGAAATTAAGAAATTAAGAA 3516
Qy 8349 AGAAGAAATTCGCGAGCAGAGAGTGAATCTTCTACAGCACAAGTAAATTAAGTAACTCCGT 8408
Db 3517 AGTTAAAAAAGAGTAAAGAAAAAGAAATTAATTAAGAAATTAAGAAATTAAGTATAGT 3576
Qy 8409 AGTTGTTTACGAAAGCAAGATTAATGAATTAATGATTAACACAAAAAATATATTTTCAGA 8468
Db 3577 ACAAGAAATTAATGAACGAAGATGTTAACGAAGAAATCCGCAAAACAAGATTAAGTAT 3636
Qy 8469 AGTCAATGCTCTGCTTAAATGATACAAAGAAAGAAAGTAAATTAAGATCTTTGCGCGT 8528
Db 3637 AGAACAAAAAGAAAAAGAAAGTAAAGAAAAAGAAAGAAAGTAAAGAAAAAGAAAGA 3696
Qy 8529 AGCGGTGTGATGACACAAAGAAACAAGCAATTTACGAGATCAACAAAGTTAACTTC 8588
Db 3697 AGTAAAGAAAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAG 3756
Qy 8589 TACAACTGTAAATGAGAAACGTATCTCACTTGTGCAAAAGCTTTGGCTTAATATGA 8648
Db 3757 ACAGAAAAAGAAAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGAAAGA 3816
Qy 8649 AAATTAATGAATTAATGAAGAACTGGAAGGCTTGTAGCGAGCGGAACAGCAGCCGT 8708
Db 3817 TAAAGTATGAGACAAAGAAATTAATTAAGAAATTAAGAAAGTAAAGAAAGTAAAGAAAGT 3876
Qy 8709 TGAATATTAACAAGATTAACA-----GAGCAATTTGTTGACAGAAATTTGGA 8759
Db 3877 AAAAAAAGAAATTAATTAAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3936
Qy 8760 AATTGAGATTAATTAAGAAAGATTCGAAAGATTAATGATTTGAAGTCAACGAGAGA 8819
Db 3937 CGAAGATGTTAAACGAAAAAGATACCGCAAAACAAGATTAAGGTATGAACAAAGAAAAAGA 3996
Qy 8820 CGGAACCAAGAGGCTCTTGTGGAAGAAAGTATTTCTGTGAAATTAACAATTTCCAG 8879
Db 3997 AAAAGAAAGATTAAGAAAGTAAAGAAAAAGATTAAGTAAACAAAGATTAAGTAAAG 4056
Qy 8880 GGAACCAATATCATCTGAAGATTAAGCAAGATTTGTTGAACCGGAAGGTAAATGT 8939
Db 4057 ACAGAGATTAATTAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGAAAGAA 4116
Qy 8940 AGATGCTTTGAATGAATGATTAATGATCTCAAGAAAAAGTGTGCTAATGTGAT 8999

Db	4117	TAATTAATAATGAAAAACAAGATAATGTAGTACAAAGAAATAATGACGAAGATGTTAA	4176
Qy	9000	TGCTATTGGAATGTTGATGTAAATATGTAAGAAAAATGTAAGACCAAAATCGG	9059
Db	4177	CGAAAAAGATACCGAAACCAAGATAAATGATGAGAAAGAAAGTAATAATGAAAGAGT	4236
Qy	9060	AAAGCATCTTTGTGAGAAACCTACGAAAAACAAGATATCAAGCATTTTAAACAGGCAA	9119
Db	4237	AAAAGAAAGATTAATAAAAGAGTAACAAAGAGTTTAAACAAAGATTAACAGAGAA	4296
Qy	9120	AGTAATATCTCTTGGAAG 9139	
Db	4297	TAGAAAAATGAAAAAG 4316	
RESULT 18			
LOCUS	A58933	5361 bp	DNA
DEFINITION	Sequence 2 from Patent WO9641877.		linear
ACCESSION	A58933		
VERSION	A58933.1		
KEYWORDS	GI:3714407		
SOURCE	unidentified		
ORGANISM	unclassified		
REFERENCE	1		
AUTHORS	Drullhe, P. and Daubersies, P.		
TITLE	MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES		
JOURNAL	Patent: WO 9641877-A 2 27-DEC-1996; PASTEUR INSTITUT (FR) AU 6309666 970109 Other publication FR 2735478 961220. Location/Qualifiers		
COMMENT			
FEATURES			
source	1..5361		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
	1..5361		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAA03511.1"		
	/db_xref="GI:3714408"		
	/translation="MTNSNYKSNKNTYNNNEQITTTFNRTNMPICKCHREKINK YFPLKILCTLLIWAQYDNNSDINKSKNTYVDKLNKFNLSGQSVNGLAS EEVEKILDLLEEQNTLTESVDNDKLEAEEDIKENILSNEPEKNTIDMLNNG QNSEKQESYENVQYSDLEFNELNLSVDNGVEKENILLESQVNDIPNLYKSVQOE OOHNEKKEESVEENDESVENVEENVEENDQGSVASVESIASVDESIDSIE ENVAAPVEEIVAPSVESVAPSVESVEENVEESVAENVEESVAENVEESVAENVEES VAENVEEIVAPVEEIVAPVEEIVAPSVESVAPSVESVAENVEESVAENVEESVA ENVEESVAENVEESVAENVEESVAENVEEIVAPVEEIVAPVEEIVAPVEEIVAPVE EESVEENVEESVAENVEESVAENVEESVAENVEESVAENVEESVAENVEESVAENVE ESVAENVEEIVAPVEEIVAPVEEIVAPSVESVAPSVESVEEVEESVAENVEESVAENVE VAENVEESVAENVEESVAENVEEIVAPVEEIVAPVEEIVAPSVESVAPSVESVE EENVEESVAENVEESVAENVEESVAENVEEIVAPVEEIVAPVEEIVAPSVESVAPSVESVE VESVEENVEESVAENVEESVAENVEESVAENVEESVAENVEEIVAPVEEIVAPVE EIVAPSVESVAPSVESVEEVAENVAITNLSNDLSNLGGIETEEIKDLSLNEIEVEKENV VAPSVESVAPSVESVEEVAENVAITNLSNDLSNLGGIETEEIKDLSLNEIEVEKENV TILLENVEETTAESVTFPNILIEIQENITNDTIEKLEBLENVLSALENTOSE EKEEYIVDIEVEKENVATTLIEVPOAEKSNATITIEPNIENNAVEENVAENLE KANEIVFNTVLNVEETVEISGSELENEMDKAFSEITDNKQIOENILTMFRSIE TSIVIOSEKQVDFNLNENVASIINDIENMKELNKLENTISTEGVEETVEHQNVY VVVDVAMKQDFGLINEAGGLKEMFPNEDVFKSESDIVTVEEIDEBQKQVEKET VGIIEEMENIIVDLEEKEDLTDKMDAVEESIEISDSKEETESIKDKEKQVSI BEYQNDMESVEKVELKQMEBELKMDAVEINDITSIDISKEETESIKDKEKQVSI KKELEKALSDESKETIIDAQDTLEKVEIEBHDTITTLQVVELKQVEDKIKVEDL KQLEEDITLKVEIKETLEESIELEDYKELTITIEDILIEKKELIKDHPKPEPEAEIK DLEADITLKVESLSLEVEEKLEVEHLEKSEVHTISGDATIKLEEDLEBVDLGS ILDMLGDMELGDMKESLEDVTTKGERVESLKDVLSSAGIDEMQTRKKAQKPK LEEVLKEVEKEPKKKTICKKVPDIKKEPEDEIVEKMDIEDIEBVEEDIEED1 EBDKVEDIDIEDIDEDIGEDKQVILVQVEKRIEKYKAKKKLEKKVEGSGALKH VDEVMKYQKIDKEDKVAESKALESKNDLVNKKONODPSKKNPKKKKVAAPFI SAVAAPASYVGFPTSLPSSCTTIASTYLLSKVDTINKNERPFYSVPIPKNL KTYLQMKKEFKSKENNNVLEVTNKAKEKNQVYTKTEKTTTVYDKNKPKKRRQK		

ORIGIN	SK"		
Query Match	1.24;	Score 116.2;	DB 6; Length 5361;
Best Local Similarity	41.74;	Pred. No. 4.9e-06;	
Matches 1747;	Conservative 0;	Mismatches 2373;	Indels 67; Gaps 14;
Qy	5126	GTTCGTGAGAAACGGAAGCTTTAGTGTAGTATTCGTGAGTTTGAAGAGTAAGTTCTTCA	5185
Db	773	GTGTGAAGAAAGTATGATCTTCAAGTGTGTGATGAAGATATGATTTCAAGTATGAAGAA	832
Qy	5186	ATGTAGATGCAAAAGATTCAAAACAAATTAATACAA--TTGCCGGAATCCAAATGAG	5242
Db	833	ATGTAGATCCCAAGTGTGAAGAAATCGTAGCTCCCAAGTGTGTGAAGAGTGCGTCCAA	892
Qy	5243	GAAGACCGCTGAGATTGAGCAACGTTGCTCATCAAAATTTGGAAGAAACAATAGTTA	5302
Db	893	GTGTGAAGAAAGTGTAGAGAAAGTGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAA	952
Qy	5303	TAGCTATGTATAAAACGTAATAATTACAGGCGAATGATCAAGATGAAAAATATCA	5362
Db	953	GTGTAGCTGAAAAGTGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGTAA	1012
Qy	5363	ATGTGATGCAAAAGATTATATCTATGACCAATATCTATGACAGTCCGAGTTGAGAGCA	5422
Db	1013	ATGTGAAGAAATGCTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAA	1072
Qy	5423	AAGAGCCTCTGTGCAAGAGCTTCTGCAAGTACTGACTTGAATTAAGACAGTTTCTTC	5482
Db	1073	TTGTAGCTCCAAAGTGTGTAGAAAGTGTGCGCTCCAAAGTGTGMAAGAAAGTGTGAAGAA	1132
Qy	5483	ATGTGATCAAACTGATATTTGACAAAGATTTAGAGAGAAATTAATGAAATTAAGAA	5542
Db	1133	ATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAA	1192
Qy	5543	AGGCAATGTTATATGTTAGCTGAAAATACGAGTCAAGTGTCACAATGCGACAGTGC	5602
Db	1193	GTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAA	1252
Qy	5603	TTTCCGAGCAAGTGAAGAGCTGCAAGTGAAGAGCTGCAAGTGAATTAATAATTAC	5662
Db	1253	ATGTGAAGAAATGCTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAA	1312
Qy	5663	AAAAATCTTGTGACATATTAATAATATGATCAAAATGTAAGAAATGTTGTTGTTAA	5722
Db	1313	TTGTAGCTCCAAAGTGTGTGAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA	1372
Qy	5723	GC---AAATCTCATCTATTAATAACAATTGGAATTGAGCTGAGTTGAGCTGAG	5779
Db	1373	ATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAA	1432
Qy	5780	GAGCTGAGTGAAGAGTGTGAGCAAGTGAATTAAGATTGTAAATTAATGATAGCAAA	5839
Db	1433	GTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAA	1492
Qy	5840	TAAATCATGCAAAATCACTGCGAAGGGAATGTGGAGTTATTCAGAGTCTGATGCG	5899
Db	1493	ATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAA	1552
Qy	5900	TAAATGCTAAATTATGAGGAGCAAGTGTGAGTGGCCGCTGAGCAATAGAGCTCA	5959
Db	1553	TGTAGCTCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAA	1612
Qy	5960	CCAGTGTGAATGAATTACAGATCTACAAAAGCATATGTAAAGATTCTACAGTATG	6019
Db	1613	GTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTAGAGAAATGTTGAAGAA	1672
Qy	6020	CTAAGAAAGAAAGATGATTAATTAATTAATCAACAGGCAAGTGAAGTAAAGTGTAGAT	6079
Db	1673	GTGTAGCTAATAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAA	1732
Qy	6080	AAGTATTCAAAATCTTAATTAATTAATGAAGAGCTTATCAAAAAGAAATTAAGTAAT	6139

Db 1733 ATGTTGAAAGAAAGTAGTACT-----GAAATGTTGAAGAAATCGTAGCTCCAACTGTTG 1786
Qy 6140 AAAAGAGATTTGTACCAATAGTACAGTACTCATCTTAAATCTTATTTGCGAAATG 6199
Db 1787 AAGAAATGTAAGTCACTGTTGAAGAAATTTGTAGCTCCAAAGTGTGTA--GAAAGTG 1843
Qy 6200 CCGGTGTTGACAGCAAGCCGAGAGTGGCAAGAACTGTTAATATCAACAAGTTTATGAG 6259
Db 1844 TGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATG 1903
Qy 6260 AAAAGAAAGCTCTTGTAGAAATTTCTATATTAAATGCAAAACATTAATCTGTAAATCAG 6319
Db 1904 TTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCG 1963
Qy 6320 GAGATTACAGCAATTCAGATCGAGATAGTAGTCTGTGTGTGTGTGGAATGTAGAG 6379
Db 1964 TAGCTCCAACTGTTGAAGAAATGTAAGCTCCAACTGTTGAAGAAATTTGTAGCTCCAAAGTG 2023
Qy 6380 TAGAGAGCTTCTTGATACCAATATTATTAAG--AAATACCAAGCAAGAGTTGGA 6436
Db 2024 TTGTGAAGAAAGTGTAGCTCCAACTGTTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTG 2083
Qy 6437 AAACACTCAATGTTCTGATGAAGTTTCGAGAAAGAAAGCTGAAATTCACAGCAATTTCAAGC 6496
Db 2084 TAGCTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATG 2143
Qy 6497 AAGGAATTTCTCTTTTGAAGTCCGAGTCCAGCAAGCCGAGGTAGAGCCGAGAGTGCGAG 6556
Db 2144 TTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATGTAAGCTCCAACTGTTGAAGAAATCG 2203
Qy 6557 GAAACGTTTCCGTAAATCAATTTGCAAGAAAGCAAGAAAGTAGATGTGAAGAAAGCA 6616
Db 2204 TAGCTCCAACTGTTGAAGAAATTTGTAGCTCCAAAGTGTGTAGAAAGTGTGCTCCAAAGTG 2263
Qy 6617 TTTTGTGTAAGAAAGTGTGATTAACAGCAAAAGCTTATAGTCTGTGTGCAATTTGAAATG 6676
Db 2264 TTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTG 2323
Qy 6677 CCGCAGTCCGAGTGTGCTGCAAAAGAGTGTGAATTTGAGCAGCAGTGGCAGTTACCAAG 6736
Db 2324 TAGCTGAAATTTGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTCCAACTG 2383
Qy 6737 ATGAATCAAAACGAGAGCAGAGTAGTGAATTTCTAAAT--ATGACTCGAAACAAAGT 6793
Db 2384 TTGAAGAAATTTGTAGCTCCAAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGTG 2443
Qy 6794 TAGATGTAATAGCAAGAAATGAGATTAATCAGAGTCTGGAATCGGTTCAAGCCGAGCTG 6853
Db 2444 TTGCTGAAAGCTGTGCAACAAATTTATCAGCAATCTTTTAAGTAATTTATTAAGTGTGTA 2503
Qy 6854 GAAATCTTGCAAGCCGAGATATCGAGTGTGTTCTGTCAATTAATTTGCAATTAAGTATG 6913
Db 2504 TCGAAACTGAGAAATTAAGACAGTATATTAAATGAGATAGAGAGTAAGAAAGAAATG 2563
Qy 6914 AAACAGATATCGATCATAGTACTTTTACACTCTTCTACTGATGTAAATGTAAGAGCTCTTA 6973
Db 2564 TAGTCACCAATATCTAGAAAAAGTGAAGAAAGTGAAGAAAGTGAAGTGTAACTATCTTTTA 2623
Qy 6974 ATTAATTTTTCGAATTTCTTGACAGCCGAGTGAAGAGCCGAGCTCTTGACAGCATTTACCG 7033
Db 2624 GTTAACATATTAGAGAGATACAGAAATATCTATTATGATATCTAATAGAGAAAT 2683
Qy 7034 GAGTGGTTTCTGTAACTATTAATAGTCTGTGATAGTACGAGTTCACATTAATCTG 7093
Db 2684 TGAAGAACTCCAGAAATGTAATTAAGTCCGCTTTAGAAATATCCCAAAGTGAAGAG 2743
Qy 7094 ATTGACTTCGATCAGAGA---AAAGTAATGTAAACGCAAAAGAGAAATTAATTA 7150
Db 2744 AAAAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2803
Qy 7151 AGCAAAACAGCAAGAAATGAGAAATGAGAGAGCAGCAATCGAGCCAAATGCTTGTGTA 7210
Db 2804 TAGAAACTGTGAAACAGCAGAGAAAGAAAGCGCAATTAACAATTAACGAAATATTTGAAA 2863

Qy 7211 ATTAATTTTGAACAGCTGTAGAGATAGAAATTAATTTCTGAAGAAAGCAAGAGTTT 7270
Db 2864 ATTTAGAAAGAAATGCAATAGAAAGTATGAAATGTTGCAAGAGATTTTGAAGAAATTTAA 2923
Qy 7271 TAAAACTTTAGAGAAATTAACAAAGCAAGATTAAGAAATTAATTAATTAATTAATTAATTA 7330
Db 2924 ACGAAACTGTATTTTAATCTGATTAAGTAAAGTAGAGAAACAGTAAATTTAGCCGAG 2983
Qy 7331 AAATCTTCAATTCAGAGGTATTTCTACAGAAAGTACTCTGTAAACCGATAGAGAG 7390
Db 2984 AAAGTTTGAAGAAACATGAATGATTAAGCAATTTTGTAGAAATTTTAATTAATTAATTA 3043
Qy 7391 ATACTCAGGAGAAAGAAATTAAGCAATTTGAAGCTTCTGATATTAATTTGAAATTAATG 7450
Db 3044 AAGGAATCAAGAAATTTATTAACAGTATG-----TTTGAAGTATTAAGAAACCAAGTA 3097
Qy 7451 TAGATATTACACAGAGACAAAGATTAATCACTTCTACTGTGTGTGTTGGAACTGAG 7510
Db 3098 TAGTATTCCAATCAGAAAGAAAGTGTGATTTGAATGAAGAAATGTGTGATTCGATTTTATG 3157
Qy 7511 GTCTTGCTTCGCAATCAGGAACAGTGGAGTTACAAATTAATTAAGAAATTCGAGATTA 7570
Db 3158 ATAAATATGAAGAAATATGAAGAAAGTGTATTAATTAATTAATTAATTAATTAATTAATTA 3217
Qy 7571 CTGTTGAAATTTCTTTTGTGAAGCAGCTGAAAGAAAGTAAATGTTAGATCGATAT----- 7625
Db 3218 AAGGTGTTCAAGAAACGTGTAATGTAACATGTGAACAAATATTAATGTGATGTGATG 3277
Qy 7626 ---TACAGAAATGTTGCTTTAACAGATATCAAGGCTCTGTAGAGACATTTGGAAATG 7681
Db 3278 TTCCGTGATGAAGATCAATTTTGAAGAAATTAATTAAGAGCAGAGGAGTGAAGAA 3337
Qy 7682 GAGTGCTATGCAAGATTAATTTCTAATGGAAGTCAATATGATTAATTAATTAATTTCTA 7741
Db 3338 TGTTTTAAATTTGAAAGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3397
Qy 7742 AGCTATTGAGAAATTTATGATGTTATTTGAAGAAATTCGAAATTTGAGAGCCGAG 7801
Db 3398 TTAAGAGTAACCGGTTCAAAAAGAGTGAAGAAAGAAACGTTAGTATTTGAAGAA 3457
Qy 7802 CAAGAGATTAACCGTAGAGCCGAGTGTGCGGAGCCATTAATCTCAAGCAAGAAATG 7861
Db 3458 TGAAGAAATATTTGATGATGATTAAGAGAAAGAAAGAAAGATTTAACACAGAAAGTA 3517
Qy 7862 AAATCAATTCAGAGTTGAATTTGAGAGATTTTCAATGAAGAAATTAAGATTAATTA 7921
Db 3518 TAGATGAGTGAAGAAATCCATAGAAATATCTCAGATTTCAAGAGAAAGCTGATCTA 3577
Qy 7922 GCCCTTCTAAAGAAATTTGAAGAGAAATCAATGTCAAAGTGGAAAGAAAGCAAGATGA 7981
Db 3578 TTAAGAGTAAGAAAGAAAGATGTTTCACTAGTGTGTAAGAAAGTTCACAGATGATTAAG 3637
Qy 7982 CTGCTGAATCTCAAGAGCT-----TGTGAGAGCAGTGAAGAGGCGCAG 8027
Db 3638 ATGAAGAGTGTGAAGAAATTTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3697
Qy 8028 AATTAATTTCCAGAGAAAGAGCCGGAAGCTCTTATTTGAAGTATTAATTAATTAATTAATTA 8087
Db 3698 CTGTTGAATTAATGACATTTCTAGCAACCTTATTAAGAAAGCTCAAGAGTTAAATGAAG 3757
Qy 8088 AAGAGATATTTTCTCATGAGATTAATGTAATTAAGCAACATTAATTAATTAATTAATTAATTA 8147
Db 3758 TAGAGCAGATTTAATTAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 3815
Qy 8148 AGCAGTTTCTAAGCAGTGAAGGTTCTGTATTTGAGAGAGTGTGAGTCAACCAAGCAG 8207
Db 3816 AGAAGATTTCTAAGAAATTAATAGATGCAAAAGATGATACATTAGAAAGAGTTATTAAGAAAG 3875
Qy 8208 AGCTACTGCTGAGGTAAATTAATGTAAGTGAAGTGAAGAAATTTGTTCAAGAAACAA 8267
Db 3876 GGAACATATTAACAGAGAGAGTGTGATGAAAGTGTGAATTAAGAAAGTGTGAAGAAAG 3935

```

QY 8268 TCGATTGATGCAATTTCTAAAGTAGAAGGTTGGATGAGATTAAGTAAGTCTGTAATC 8327
Db 3936 CAAAGTCGAAAAGATCTGATTTTAAAGATCTTGAGAGAGTATATTTAAAG----- 3988
QY 8328 TTCTGATGATCAGGAAATGAGAGAAATGCCGAGCAGGAGTGAATATCTTACAGC 8387
Db 3989 --AAGTAAAGAAATCAAGAAAGCTTGAAAGTGAATTTAGAAAGTTATTAAGAAATTAA 4046
QY 8388 ACAAGTAATCTGAATCCGATGCTTGTTCAGAAAGCAAGATTTAGAAATTAATGATTA 8447
Db 4047 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATAGAAAAAGATCATTTGAAAA 4106
QY 8448 CACAAAAATATATTTTCGAAGTCAATGCTCTGCTTAAATGATACAAAGATGAAGC 8507
Db 4107 ATTCGAGAGAGAGCTGAAGAAATATAAGATCTTGAGCAGATATATTAAGAGATATC 4166
QY 8508 GAATATGAAATCTTTAGCCGTAGCCGATGTCATGACAGAGAACAAACAAAGCATTTAC 8567
Db 4167 TTCATTAGAAAGTGAAGAAAGAAAAAATTAGAGAGATACAGAAATTAAGAGAGAGT 4226
QY 8568 GAGATCAACAAAGTTAACTTCAACATGTAATGAGAGAAACGTAATCTCAACTTCGTC 8627
Db 4227 AGAACATATATTAAGTGTGATGCGCATATTAAGGTTTGAAGAGATGATTTAGAA 4286
QY 8628 AAAAGCTTGGCTTAAAAATGAAAAATTAAGAAATGTAAGAACTGAGAGACCTTACT 8687
Db 4287 AATGATGTA-TTTTAAAGAGAGATATATGACATGTTAAAGGAGATATGA-ATTAG 4343
QY 8688 CGAGCGGAAACAGCAGCCGTTGAAAAATTATACAAAGATTACTACAGAGCATTCGTTGC 8747
Db 4344 GGAATGATTAAGAGAAAGTTAGAAATGTAACACAAACCTTGAGAAAGAGTTGAATC 4403
QY 8748 AGGAAATTTGGAAATTTGAGATTAATTAAGAAAGATGCAAGATTAATGATTTAAG 8807
Db 4404 CTTAAAGATGTTTATCTAGTCACTTAGCATGATGTAAGAAACAATGAAAAACAAGAA 4463
QY 8808 AGTCACGAGAGAGAAACCAAGAGAGTCTGTGCGAAAGAAATGTAATTTCTGTGAAAA 8867
Db 4464 AAAAGCTCAAGACCTAGTTGAGAGAGATATATTAAGAAAGAGTTTAAAGAGAAC 4523
QY 8868 TCAATTTTCAGGGGAAACAAATCATCATGAAATGAAAGCAGAAATTTGTCGACCG 8927
Db 4524 AAAGAAAAAATTAACAAAAAGAAAGTAAAGTTGATTTAAGATTAAGAACCAAAAGA 4583
QY 8928 AAGTGAATGTAAGTCTTTGATGAATGAATGTAAGTCTTAACAAGAAAAAGTGTG 8987
Db 4584 TGAATATGTAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 4643
QY 8988 CTATGCTGAATTTGTAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAAT 9047
Db 4644 AGAAGAGATATGAGAGAGATTAAGTGAATTAAGATTAAGATTAAGATTAAGATTA 4703
QY 9048 AGCAAAATTCGAAAGCATGCTATTGAGAACTACGCAAAACAAAGATATCAAGATT 9107
Db 4704 AGGTGAAGACAAAGATGAATTAAGATTTAATTAAGTCCAAAAAGAGAGAGATTA 4763
QY 9108 TACAGAGCAAAAGTAATATTTCTTGAAAAAGAGAGCAGTCAAGCTGCAAGTCAAT 9167
Db 4764 CGTTAAAGCGAAAAAGAAAAATTTAGAAAAAGTTGAAGAGGTGTTAGTGTCTTAA 4823
QY 9168 GAAATGACACATTTCCAATGAGATGATTTAAAAATTTGGCAAGACGATATGATCTTC 9227
Db 4824 AAAACACGTAAGCAAGTAATGAATATGTTCAAAAAATTTGATTAAGAAAGTTGATA 4883
QY 9228 TCAATTAATACCAAAATTCAAAAATTAATATTACTTTAGCATCA 9274
Db 4884 AGTATCTAAAGCTTTAGATCAAAAAATGATTTACTAATGTTTAA 4930

```

RESULT 19
AR131272 5361 bp DNA 1linear PAT 16-MAY-2001
LOCUS
DEFINITION Sequence 2 from patent US 6191270.

```

ACCESSION AR131272
VERSION AR131272.1 GI:14119597
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5361)
AUTHORS Drullhe P. and Dauberties P.
TITLE Malerial pre-erythrocytic stage polypeptide molecules
JOURNAL Patent: US 6191270-A 2 20-FEB-2001;
FEATURES
    source
        location/Qualifiers
            1..5361
            /organism="Unknown"
            /mol_type="unassigned DNA"

Query Match 1.2%; Score 116.2; DB 6; Length 5361;
Best Local Similarity 41.7%; Pred. No. 4.9e-06;
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

QY 5126 GTTCTGATGAACGGAAGCTTTAGTATGATTCGAGTTGAGAGATTAAGTTCTTTCA 5185
Db 773 GTGTTGAAGAAAGTATAGCTTCAAGTGTGATGAAGATTAATGATTCAGATTAAGAA 832
QY 5186 ATGTAGATGCAAAAGATCAAAATCAATTAATACAA---TTGCCGAAATGCAAAATGAG 5242
Db 833 ATGTAGCTCAACTGTTGAAGAAATCGTAGCTCCAGTGTGTAAGAAAGTGTGCTCAA 892
QY 5243 GAAAAACGCGCTGAGAGTTGAGACACAGTTGCTCATACAAATTTGAAAAACAATCAGTTA 5302
Db 893 GTGTTGAAGAAAGTATGAGAAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 952
QY 5303 TATGCTATTTGTAATAAAACAGTAATTAACACGCGCAATGATCAAGTATGAAAAATATCA 5362
Db 953 GTGTAGCTGAATATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGA 1012
QY 5363 ATGTAGCTGAAGAAAGTATTAATGATGACCAATATATAGCAGTTCGAGTTGAGAGCA 5422
Db 1013 ATGTGAAGAAATCTGATGCTCAACTGTTGAAGAAATCGTAGCTCCACTGTTGAAGAA 1072
QY 5423 AAGGAGCTCTGTGCAAGAGCTTTCGAAGTACTACCTTGAAATAGACAGTTTCTTCTC 5482
Db 1073 TTGTAGCTCAAGTGTGTAAGAAAGTGTGCTCAAGTGTGAAGAAAGTGTAGAGAA 1132
QY 5483 ATGTGATCAAACTGATATTGACAAAGATTAGAGAGAAAAATATGAAAAATAGAGAA 5542
Db 1133 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAA 1192
QY 5543 AGGCAATGTTATGTTCTAGCTGAAGAAATACGAGTCAAGTGTGCAAAATGCGACAGTGC 5602
Db 1193 GTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGA 1252
QY 5603 TTTCCGAGCAAGTGAACAAGCTGCAAGTGAAGCTGAGTAGCAAGTTAAATTAATTAAC 5662
Db 1253 ATGTGAAGAAATCGTAGCTCAACTGTTGAAGAAATCGTAGCTCAACTGTTGAAGAA 1312
QY 5663 AAAATATCTTCTGCAATATTAATAATATGATCTCAAAATGTAAGAAATGCTTTGTA 5722
Db 1313 TTGTAGCTCAAGTGTGTAAGAAAGTGTGCTCAAGTGTGTAAGAAAGTGTAGAGAA 1372
QY 5723 GC---AAATCTATCATATTAATAAACAATGGAATTTGAGCTGAGTGTGAGCTGAG 5779
Db 1373 ATGTGAAGAAAGTGTAGCTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAA 1432
QY 5780 GAGCTGAGTGAACAGGTTCTGTAGCAGTAATTAAGTTGTAATTAATACATAGCAAGAT 5839
Db 1433 GTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGA 1492
QY 5840 TAAATCATGCAAAATCACTGCGAAGGAAATGTCCGAGTTATTAACAGACTCGATGCGG 5899
Db 1493 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAA 1552
QY 5900 TATATCTATTAATGAGAAAGTGTCTGAGAGTGCCTGTCAGCAATATGAGACCTCA 5959

```

Db 1553 TCGTGGCTCCAACTGTGAAAGATCGTAGCTCCAACTGTTGAAGAAATTGAGCTCCAA 1612
 Oy 5960 CCAAGTGTGATGAATTACAGAGTCTACAAAGCATATGTAAAGATCTACAGTATG 6019
 Db 1613 GTGTGTGTAAGAAAGTGGCTCCAGTGTGAAAGAAAGTGTGAAGAAATGTTGAAGAA 1672
 Oy 6020 CTAAAGAAAGAAAGATGATTAATTACTACTCAAGGCAAGTAGATAAAGTGTAGATA 6079
 Db 1673 GTGTGCTGTAATGTTGAAGAAAGTGTGTAAGAAATGTTGAAGAAAGTGTAGCTGAAA 1732
 Oy 6080 AAGTATTCAAAATCTTAATATTAACGAAGACTTATCAAAAAAGAAAAATAGTAATA 6139
 Db 1733 AATGTGAAGAAAGTGTAGCT-----GAAAAATGTGAAGAAATCGTAGCTCCAACTGTG 1786
 Oy 6140 AAAAGGATTTGTACCAATAGTACAGTACTCATACTTAATTAATCTTTATTTGCAAAATG 6199
 Db 1787 AAGAAATCGTAGCTCCAACTGTTGAAGAAATGTGTAGCTCCAACTGTGTA---GAAAGTG 1843
 Oy 6200 CCGCTGTTCAGAGACAAAGCCGAGAGTGTGAGAACTGTTAATATCAACAAGTTTATGAG 6259
 Db 1844 TGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAAATG 1903
 Oy 6260 AAACAGAAAGCTCTGTGAAGAAATTCATATTAATGCAAAACATTAATCTGTAAATCAG 6319
 Db 1904 TTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCG 1963
 Oy 6320 GAGATTACAGAAATTCATCGAGTAGTAGTCTGTGTGTGTGTGTGAAATGTAGAGG 6379
 Db 1964 TAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTGTAGCTCCAAAGTG 2023
 Oy 6380 TAGAGACTTCTTCTGATACCAATATTTATTAAG---AAATACCAAGACAAAGTTGAA 6436
 Db 2024 TTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTG 2083
 Oy 6437 AAACATCAATGTCTGATGAAGGTTTCGGAGAAAGAGCTGAATTTACAGACATTTCAAC 6496
 Db 2084 TAGCTGAAGAAATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATG 2143
 Oy 6497 AAGGAATTTCTCTTTGTGAGTCCGAGTCCAGACAGCCGGGTAGAGCCGAGTGTGAG 6556
 Db 2144 TTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCG 2203
 Oy 6557 GAACGGTTTCCGTAAATCAATTTGCGAAGAAAGACGAAAGTAGATGTGAAGAAAGCAAGA 6616
 Db 2204 TAGCTCCAACTGTTGAAGAAATGTGTAGCTCCAAAGTGTGAAGAAAGTGTGCTCCAAAGTG 2263
 Oy 6617 TTTTGTGTAAGAAAGTGTGATTTACAGCAAAACGTTATAGTCTGTGTGCAATTTGAAATG 6676
 Db 2264 TTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTG 2323
 Oy 6677 CCGAGTCCGAGTGTGCAAAAGAGCTGTGAATTTGAGACAGACAGTGTGCAATTCACAAAG 6736
 Db 2324 TAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTCCAACTG 2383
 Oy 6737 ATGAATCAAAACAGAGACAGAGTAGTAAATTTCTAAATTT---ATGACTCGAAACAAAGT 6793
 Db 2384 TTGAAGAAATTTAGCTCCAAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGTG 2443
 Oy 6794 TAGATTAATAGCAAGAAATGAGATAAATCAGAGTCTGGAATCGGTTCAAGCCGAGCTG 6853
 Db 2444 TTGCTGAAGAAAGTGTGCAAAATTTATCAGACATCTTTTAAGTATTTATTAAGTGTGATA 2503
 Oy 6854 GAATCTTGACCGAGAGTATCTGAGTGTCTGTCAATAATTTGCAATTAAGTGTG 6913
 Db 2504 TCGAAACTAGAGAAATTAAGACAGATATTAATATGATTAAGAGAAAGTGAAGAAATG 2563
 Oy 6914 AAACAGATATCATAGTATTAATCTTACACTCTTCTAGTATTAATGTAAAGCTCTTA 6973
 Db 2564 TAGTACCAACAATACAGAAAGCTGAAGAAAGTGTAGCTGAAAGTGTATTAATCTTTTA 2623
 Oy 6974 ATTAATTTTCGAATTCCTTGACAGCCGCTGTGAGAGCCGAGCTCTTGACAGCATTAACG 7033

Db 2624 GTAAATATTAGAGAGATACAGAAATACTATTACTAATGATATCATAGAGAAAAAT 2683
 Oy 7034 GAGTGTGTTCTGTTAACACATATAATAGTCTGTGATAGTGTGAGATTCAATTAATCTG 7093
 Db 2684 TGAAGAACTCCAGAGAAATGTATTTAGTGTGCGCTTTTGAAGAAATACCAAGTGAAGAG 2743
 Oy 7094 ATTTGACTTCGGTACGAGA---AAAAGTAAATGTACCGCAAAAGAGAAAAATATTA 7150
 Db 2744 AAAAGAAAGAAATATAGATTAATTTGAAGAAAGTGAAGAAAGAGTGTGCTACCACTTTAA 2803
 Oy 7151 AGCAAAACGACGAAATCCAGGAATCCGAGAGACGAATGGAAGCCAAATGCTTTGTGTA 7210
 Db 2804 TAGAACTGTGTGAACAGGCAAGAAAGAGCCGAATTAATCAATTCGAAATATTTGAAA 2863
 Oy 7211 ATTAATTTTGAACAGCTGTAGAGATAGAAAAATTTCTGAAGAAAGAAACAGAAAGTTT 7270
 Db 2864 ATTTAGAGAAATATGCAATGAAAGTAAAGAAATGTTGTGCAAGAAATTTAGAGAAATTA 2923
 Oy 7271 TAAAACTTTAGACGAAGTTAAACAAAGACAAAGTAAAAAGTAAATGATCTACGAAA 7330
 Db 2924 ACGAAACTGTATTTAATCTGTATTAAGTAAAGTGAAGAAACGTAAGAAATTTAGCCGAG 2983
 Oy 7331 AAATCTTCAATCAGACAGTATTTCTACAGAAAGTACTCTGTGTAAGCCGATAGAGAG 7390
 Db 2984 AAAGTTTGAAGAAACAAATGAATGAATGAATGAAATTTTGTAGAAATATTTATATATGTA 3043
 Oy 7391 ATACTCAGGAGAAAGAAATTAAGCATTTGTGAAGACTCTGTATTTATGGAAGAAATG 7450
 Db 3044 AAGGAATCAAGAAATTTATTAACAGATAG-----TTTGAAGTATTAAGAACCGATA 3097
 Oy 7451 TAGATATTACAACAGAGACAAAGATTAATATCATCTTACTGTGTGTGTGGAACCTGAG 7510
 Db 3098 TAGTATTCCAATCAGAGAAAGAGTGTGAAATGAAAGATGTGTGATTTGCTGATTTTAA 3157
 Oy 7511 GTCTTGCTCCGCACTCAGGAACAGTGTGCAATTAATTAAGAAATTTCCGAGTTA 7570
 Db 3158 ATAAATAGAAATATAGAAAGAGTTTATTAATTAATTAAGAAATTTTCAAGTACTG 3217
 Oy 7571 CTGTGAAGAAATTTCTTTGTGAAGACAGCTGAAAGAAATGTTAGATCGATAT----- 7625
 Db 3218 AAGGTTCAGAAAGCTGTAACTGAACATGTGAAACAAATGTATATGTGATGTGATG 3277
 Oy 7626 ---TACGAAATGTGCTTTAACAGCATATCAAGTCTGTAGAGACATTTGGAAATAG 7681
 Db 3278 TTCCGTATAGAAAGATCAATTTTGAAGAAATTAATTAAGAGGACAGAGGTTGAAGAAA 3337
 Oy 7682 GAGCTGCTATGCAAGATTAATTTCTAATGGAAGTCAAAATATCATGTTTAAAAATTTCTA 7741
 Db 3338 GTTTTATTAATTTGAAAGATGTATTTAAAGTGAAGTGAATTAATCTGTAGAGAAA 3397
 Oy 7742 AGCTATTAGAAAAAATATGATGTTATTGTAAAAAGATTAATCGAATTTGAGAGCGAAG 7801
 Db 3398 TTAAGATGAACCGGTTCAAAAAGAGGTGAAAAAGAAACGTGTAGTATTTATGAAGAAA 3457
 Oy 7802 CAAGAGATTAACCGTAGAGCCGTAAGCTGTCCGAGCCATTAATCTCAAAAGCAAGAAATG 7861
 Db 3458 TGAAGAAATATTTAGATGATTAAGAGAAAGAAAGAAAGATTTAACAACAAGATGA 3517
 Oy 7862 AAATGAATTCAGAGTTGAAATTTGAGAGAGATTTTCAATGAAGAAATAGAGTAACTA 7921
 Db 3518 TAGATGAGTGAAGAAATCCATAGAAATATCTTCAGATTTCAAGAAAGAAACGTGAATTA 3577
 Oy 7922 GCCCTTCAAGGAATTTGAAGAGAAATCAATGTCAAGTGAAGTGAAGAAAGCAAGATGA 7981
 Db 3578 TTAAGATTAAGAAAGAAAGATGTTTCACTAGTGTGTTGAAGAAAGTTCAAGACATGATAAG 3637
 Oy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGACAGTGAAGAGGACAG 8027
 Db 3638 ATGAAGAGTGTGAGAAAGTTTGAATTTGAAGAAATATGAAGAGAGATTAATGAAGATG 3697
 Oy 8028 AATTAATTTCCAGAGCAAGAGATGCCGAAAGCTCTTATTTGAAGATTAATCAAAAATCCG 8087
 Db 3698 CTGTGAATTAATGACATTACTAGCAAACTTATTTGAAGAAACCTCAAGAGTTAAATGAAG 3757

Db		8088	AAGAAGATTTTTCATGCACATAAAGGAATTGGAAGCAACAATTAAATGAAAGTAAC	8147
Oy		8088	AAGAAGATTTTTCATGCACATAAAGGAATTGGAAGCAACAATTAAATGAAAGTAAC	8147
Db		3758	TAGAAGCAGATTTATATATAAAGATGTGGAAAAATTTAAAGAAATTGAAAAGAAC--ATTATC	3815
Oy		8148	AGCAGTTTCTAAAGCACATAACAGGTTCTGTATTGGAGAGAGTTGAGTCCACCAAGCAGA	8207
Db		3816	AGAAGATTCTTAAGAAATTAATAGATGCAAAAAGATGATACATTGAAAAAAGTTTGTAGA	3875
Oy		8208	AGTACTGCTGCAGGTAAACTATGTAGAAAGTTGAGGAAGAAATTGTTCAAGACAA	8267
Db		3876	GGAACATGATATATACGACGAGCTTGATGAAAGTTGTGAATTAAAGATGTCCAGAGAGA	3935
Oy		8268	TCGATTGAATGCCATTTCTAAAGTAGAAGGTTTGGATGGAAGATTAAGTAATCTGCTAATC	8327
Db		3936	CMAAGATGAAAAAATATCTGATTTTAAAAAGATCTTTGAAAGAAAGATATATTTAAAG-----	3988
Oy		8328	TTCTGTAGTATCAGGAATGAGAGGAATTGCCGAGCAGAGATGAAATTACTTCAAGC	8387
Db		3989	--AAGTAAAAAGAAATCAAAGACTTGAAGTGAATTTTGAAKATTTATTAAGATTAA	4046
Oy		8388	ACAAAGTAATCTGAAATCCGTAAGTTCGTTTACGAACCAAGATTATGAAAAATATGATTA	8447
Db		4047	AACATTTGAACACAGATATTTTGAAGAAGAAAAAGAAATGAAAAAAGATCATTTTGA	4106
Oy		8448	CACAAAAAATATATTTTCAGAACTCAATGCTCTTGTCTTAAATGATATCAAAAGATGAGC	8507
Db		4107	ATTGGAAGAAAGAACTGAAGAAATATAAGATCTTGAAGCAGATATATTTAAAGAAAGTATC	4166
Oy		8508	GAAATATGAATCTTTAGCGGTAGCCGGTGTGCATGCACAGAAACAACAAAGCATTTAC	8567
Db		4167	TTCAATTGAAAGTTAGAAAGAAAAAAATTTAGAAAGAGTCAACGAATTTAAAGAAAGGT	4226
Oy		8568	GAGATCAAAACAAGTTAACTTCTCAACCTGTAAATGAGAGAAACCTATCTCACTTCGTCG	8627
Db		4227	AGAACTATTAATAAGTGTGATGCGCATATATAAGTTTGGAGAAAGATGATTTAGAGA	4286
Oy		8628	AAAAAGCTTTGGCTAAAAATGAAAAATTATGAAATGTAAANAGAACTGAGAGAGCCTTAGT	8687
Db		4287	AGTAGATGA-TTTTAAAGGAAGTATATTAGACATGTTAAAGGAGATATAGGA--ATTAGG	4343
Oy		8688	CGAAGCCGAAACAGCAGCCGTTGAAAAATTATACAAGAGTCTACAGAGCATTTGGTTGC	8747
Db		4344	GGATATGATATAGAAAGATTAGAAAGTATACAAACAACAACTTGAAABAAGAGTTGATC	4403
Oy		8748	AGGAAATTGGGAAATTGAGATTAATTAGAAACGATTGCAAGATATACGATTGAG	8807
Db		4404	CTTAAGAGATTTTTATTTATGTAGTGCATTAGCGATGAGATGAAACAAATGAAACAGAA	4463
Oy		8808	AGTCAACCGAACCGAACCAAGAGGCTTGTGCGAAAGATGTATTTCTGTGA	8867
Db		4464	AAAAGCTCAAAAGACTAAGTTGGAAGAAAGTATTTATTAABABAAGGTTAABGAAGAAC	4523
Oy		8868	TACAATTTCAAGGGGAAACAAAATCATCCATTGAAGATTAAGCCGAATTTGGAACCGG	8927
Db		4524	AAAAGAAAAAATTAACAAAAAAGAAAGTATGAGTTGATATTAAAGATTAAGAACCAAAAGA	4583
Oy		8928	AAAGTAAATGTAGTCTGTTGAATGAACTTGATGTAGATCTTACAAGAAAAATGGTGG	8987
Db		4584	TGAATATATGAAGTTGAATGAAAGATGAAAGATATAGAACAMAGATGTAGAGGAAGATAT	4643
Oy		8988	CTATGCGNGAAATTGGTATTTGGAAATGTTGATGTATAATATGTGATTAAGAAAAATGTAGA	9047
Db		4644	AGAGAAAGATATAGAAAGAGTAAAGTTGAAGATTTAGATGAAGATATATATGTAAGATAT	4703
Oy		9048	AGCCAAATTCGGANAGACATGCTATTTGTAGAAACTACTGGAACAAACAGATATCAAGCATTT	9107
Db		4704	AGGTGAACAAAGATGAAAGTTATATGATTTATATGTCCA AAAAGGAAAGCAATTTGAAA	4763
Oy		9108	TACAAGACAAAGTAAATATTTCTTGGAAAAAGAGACGCTCAGCTCAGCTGCATATATC	9167
Db		4764	GGTTTAACCGAAAAAGAAAAATTTGAAAAAAAAGTTGAAAGAGTGTATTAGTGTCTTAA	4823

[illegible]

Db 2612 TTGCTGAAAAAGTTGGCAACAAATTTATACGACATCTTTTAAGTAATTTAGGTGGTA 2671
 Qy 6854 GAATTCCTGACCCGGAATCTGGAGTGGTTTCTGCAATATATTTGCAATTAAGGTG 6913
 Db 2672 TCGAAACTGAGAAATTAAGGACAGATATTTAAATGAGATAGAGAAAGTAAAGAAATG 2731
 Qy 6914 AAAACGATATCGATCATACTTACTTACCTCTTCTACGTGATGTAATGTAAGTAAAGCTCTTA 6973
 Db 2732 TAGTCACCAATATCTAGAAAACGTAGAAAACCTACAGCTGAAAGTGAACCTACTTTTA 2791
 Qy 6974 ATAAAAATTTGCAATTCCTTGACAGCCGCTGAGAGCCGACAGTCTTGACAGAGTTTACC 7033
 Db 2792 GTTACCTATTAGAGGATACAAAGAAATACTATTACTATGATCTATAGAGAAAAAT 2851
 Qy 7034 GAGTGGTTTCTGTAACTATTAATAGTTCTGTGATAGCTGAGTTCAATTAACCTTG 7093
 Db 2852 TAGAAGAACTCCACGAAAATGTATTAAAGCCGCTTAAAGAAATACCAAGAGTGAAGAG 2911
 Qy 7094 ATTTGACTCCGTACAGAG---AAAAGTAATGTTACCGCAAAAGAGAAAAAATATTA 7150
 Db 2912 AAAAGAAAGATTAATGATGTATTTAGAAAGTAAAGAGAGGTGCTACCACTTTAA 2971
 Qy 7151 ACCAACAGCAGCAATTCGAGATTCGAGAGCAGCAATCGAGCCAAATGCTTGGTAA 7210
 Db 2972 TAGAAACTGTGAAACGCGAGAAAGAAAGCGCAATTAATTAACGAAATATTTGAAA 3031
 Qy 7211 ATTAATTTGGAACAGCTGTAGAAAGATAGAAAAATTTCTGAGAAAAAGAACAGAGTTT 7270
 Db 3032 ATTTAGAAAAAATGCGATAGAAAGTAAGAAATGTTCAGAGAAATTTTGAAGAAATTA 3091
 Qy 7271 TAAAACTTTAGACGAGTTTACAAAGAACAGTAAATAAAGTAATGATGCTACGAAA 7330
 Db 3092 ACGAAACTGTATTTAATCTGATTAAGTAATAGAGAAACAGTAGAAATTAAGCGAG 3151
 Qy 7331 AAATCTTACATCAGCAGGATTTCTACAGAGATACCTCTGTAAGAGCGATAGAGAG 7390
 Db 3152 AAAGTTTGAAGAAACATGAAATGATTAAGCATTTTCTAGTGAATTAATTTGATTAAGTAA 3211
 Qy 7391 ATACTGAGGAGAAAGGAATTTAAAGCATTGTGAAAGCTTCTGATATTTTGGAAAAATG 7450
 Db 3212 AAGGAATACAGAAATTTATTAAACGATAG-----TTTGAAGTATAGAAACAGTA 3265
 Qy 7451 TAGATTTACACAGAGACAGAAATTAATCTCTTACTGATGTTGGAACTGCGAG 7510
 Db 3266 TAGTATCCATTCAGAGAAAGAGTTGATTTGAATAAATGTGTGATTCGATTTTAG 3325
 Qy 7511 GTCTGCTCCGATCAGGAACGTGSCAGTTTCAATATATTAAAGAAATTCGAGATTA 7570
 Db 3326 ATTAATATAGAAATATGAAAGAGGTTTATTAAATTAATTAAGAAATATTTCAAGTACTG 3385
 Qy 7571 CTGTTGAAAAATCTTTTGTGAAGACAGCTGAAAAAGTAATGTTAGATCGATAT----- 7625
 Db 3386 AAGGTGTTCAAGAACTGTAACTGAACATGTAGAACAAATGTATATGTGATGTTGATG 3445
 Qy 7626 ----TACAGAAATGTTGCTTTAAACAGCATATCAAGGTCTGTAGAGCATTTGGAAATG 7681
 Db 3446 TTCTCTCTATGAAAGATCAATTTTATAGATATTTAATAGSCAGAGGGTTTAAAGAAA 3505
 Qy 7682 GAGCTGCTATGAGATTAATTTCTAATGAGAGATCAAAATATCAGATTTAAAAATCTTA 7741
 Db 3506 TGTTTTATTTTGGAGATGTATTTAAAGTGAAAGTGATGTAATTACTGTAGAGAAA 3565
 Qy 7742 AGCTATTAGAAAAAATATTGATGTATTGTAAAGATTAATGGAATTTGAGACCGAG 7801
 Db 3566 TTAAGATGAAACCGGTTCAAAAAGAGGTAGAAAAAGAACTGTAGATATTTAGAGAAA 3625
 Qy 7802 CAAAAAGATTAACCGTATGAGAGCGGTAGCTCCGAGCATATCTCAAAAGCAAAAGATG 7861
 Db 3626 TGAAGAAATATTTGATGTATTAGAGAGAAAAAGAAATTTAAACGACAAAGTGA 3685
 Qy 7862 AAATGAATTCAGAGTTGAAATTTGAGAAAGATATTTTCAATGAGAAAAATAGAGTAACTA 7921
 Db 3686 TAGATGACGTAGAAAGATTCATAGAAATATCTTCAAGTTCTAAAGAAAGAACTGAATCTA 3745

Qy 7922 GCCCTTAAAGGAATTGGAGAGAAATCAATGTCAAGTGGAAAAAGAAAAACAGATGA 7981
 Db 3746 TTAAGATTAAGAAAGAAAGATGTTTCTACTAGTGTGTAAGAAAGTTCAAGCAATGATATG 3805
 Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGAGCATAGAGGGCCAG 8027
 Db 3806 ATGAAAGTGTGAGAAAGTTTATAGATGAAATATGAAAGAGAGAGTAAATGAGATG 3865
 Qy 8028 AATTAATTTCCAGAGCAAAAGATGCCGAAGCTCTTATTGAAAGTTAGTACAAATCCG 8087
 Db 3866 CTGTTGAATTAATGACATTTACTAGCAAACTTATTGAAGAACTCAAGGTTAAATGAG 3925
 Qy 8088 AAGAGTATTTTTCATGCAAGATTAATGTAAATAGAACCAACATTAATGAAGTATAC 8147
 Db 3926 TAGAAGCAGATTTTAAAGAAATGATGAAAATTTAAAGAAATTTAGAAAAAGC--ATTATC 3983
 Qy 8148 AGCAGTTTCTAAAGCAGTAAACAGTTCTGTATTTGGAGAGATTGGAGTCAACCAAGCAGA 8207
 Db 3984 AGAAGATTTTAAGAAATTAATGATGCAAAAGATGATACATTTAGAAAAAGTATTGAGA 4043
 Qy 8208 AGCTACTGCTGAGTAAACTATGTGTAGAACTGAGAGGAAATTTGTTCAAGACAA 8267
 Db 4044 GGAACATGATTAACGAGACGTTGATGAAAGTTGATGAAATTTAAAGATGTGGAAGAGA 4103
 Qy 8268 TCGATTTGAATGCAATTTCTTAAAGTAGAAGTTTGGATGAAAGTAAAGTAACTGTAATC 8327
 Db 4104 CAAGATCGAAAAAGTATCGATTTAAAGATCTTAAGAGATTAATTTAAAG----- 4156
 Qy 8328 TTCTGTAGATCAGAAAAATGAGAGAAATTCGAGAGAGAGTGAATTAATCTTACAGC 8387
 Db 4157 --AAGTAAAGAAATCAAAAGACTTGAAGTAAATTTTGAAGATTAATTAAGAAATTA 4214
 Qy 8388 ACAAGTAACTGAAATCCGTAGTCTTTACGAAAGCAAGTTATGAATTAATGATTA 8447
 Db 4215 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATGAAAAAGATCATTTGAAAA 4274
 Qy 8448 CACAAAAAATATTTTGAAGATCAATGCTCTGTTTAAATGATACAAAGATGAAGC 8507
 Db 4275 ATTCGAGAGAGAGTGAAGAAATTAAGATTTTGAAGCATTAATTTAAAGAGTATC 4334
 Qy 8508 GAATATAGAACTTTAGCGGTAGCCGCTGTCATCAACAGAGAACCAAGCAATTTAC 8567
 Db 4335 TTCAATTAGAGTTGAAGAAAGAAAAAATTAAGAGAGTACACGAATTTAAAGAGAGT 4394
 Qy 8568 GAGATCAACAGTTAACTTTACAACTGTAAATGAGAGAAACGTATCTCAACTTCGTC 8627
 Db 4395 AGAACATTAATTAAGTGTGATGCCGATTAAGGTTTGAAGAGATGATTTAGAAA 4454
 Qy 8628 AAAAGCTTTGCTAAAAATGAAATTAATGAAATGTAAGAACTGAGAGACCTTACT 8687
 Db 4455 AGTAGATGA--TTTAAAGAGAGATATTAAGACATGTAAAGGAGATATGAG--ATTAG 4511
 Qy 8688 CGAGCGAGAAACAGCAGCCGTTGAAATTAATCAAGAGTACTACAGAGCATTTGGTTC 8747
 Db 4512 GGAATGATAGAGAAAGTTTGAAGATTAACACAAACCTTGAAGAAAGCTTGAATC 4571
 Qy 8748 AGCAAAATTTGGAAATTTGAGATTAATTTAGAAAAGATTTGCAAGAGATTAATGATTTAG 8807
 Db 4572 CTTTAAAGATGTTTTATCTAGTGCATTTAGCATGTAGTGAAGAACAAATGAAAAACAAGAA 4631
 Qy 8808 AGTCAACGAGACGGAACCAAGAGAGTCTTGTGGAAGAAATGTGATTTCTGTGAAAA 8867
 Db 4632 AAAAGCTCAAAAGCTTAAGTGAAGAGATATTTTAAAGAGAGGTTAAAGAAAGACC 4691
 Qy 8868 TACAATTTGAGGGGAAACAAATCACTTGAAGATTAAGCCAGAAATTTGTTGAACCG 8927
 Db 4692 AAAGAAAAAATTAACAAAAAAGAAATGATTTGATTTAAGATTAAGAACCAAAAGA 4751
 Qy 8928 AAGGTAAATGTAGATGCTTTGAATGAACCTTGAATGATCTTCAAGAAAAAGTGTGG 8987
 Db 4752 TGAATATGTAAGAGTTGAATGAAAGTGAAGATATTAAGAAAGAGATGTGAAGAGATAT 4811

Db 2372 TAGCTCAACTGTGAAGAAATTGTAGCTCCAGTGTGTAGAAAGTGTGCTCCAGTGC 2431
Qy 6617 TTTTGTAAAAAAGCTGAGATTACAGCAAAAGTTATAGTTCGTGTCAATTGGAATG 6676
Db 2432 TTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGC 2491
Qy 6677 CCGGATCGGAGTGTGCTGCAAAAAGAGCTGGAATTGGAGCAGCAGTGGCAATTACCAAG 6736
Db 2492 TAGCTGAAAATGTGAAGAAAGTGTAGTGAAGAAATGTTGAAGAAAGTGTAGCTCCAACTG 2551
Qy 6737 ATGAATCAAAACAGAGAGCAAGAGTGAAGAAATTTCTAAAAT---ATGACTCGAAACAAGT 6793
Db 2552 TTGAAGAAATGTAGCTCCAGTGTGAAGAAAGTGTAGCTCCAGTGTGAAGAAAGTGC 2611
Qy 6794 TAGATGTATAGCAAGAAATGAGATMAATCAGGATCTGGAATCGGTTCAAGCCGAGCTG 6853
Db 2612 TTGCTGAAAACGTTGCAACAAATTTATCAGACAACTTTTAAGTAAATTTATATGAGGATG 2671
Qy 6854 GAATTCCTGACACCGGAGATATCTGAGAGTGTCTGTCAATAATATGCAAAATAGGTAG 6913
Db 2672 TCGAAAATGAGGAAATTAAGGACAGATATTTAAATGAGATAGAGAAAGTAAAGAAATG 2731
Qy 6914 AAACAGATATGATCATAGTACTTATCAGCTCTTCTAGTATGTAATGTAAAGCTCTTA 6973
Db 2732 TGATCAACCAATCTAGAAAACGTAGAAAGAACTAGCTGAAAGTGTATCTACTTTTA 2791
Qy 6974 ATAAATTTGCAATTCCTTGAACCCGCTGAGAGACCGCAGGCTTTCAGCAGTTACCG 7033
Db 2792 GTACATATTAAGAGGATACAGAAATATCTATCTAATGATATCTATAGAGAAATAT 2851
Qy 7034 GAGTGTCTCTGTAACTATTAATAGTCTGTATAGCTGAGTTCACTATACCTG 7093
Db 2852 TGAAGAACTCCACGAAAATGTATTAAGCCGCTTGAAGAAATACCCAAAGTGAAGAG 2911
Qy 7094 ATTGACTCCGTAAGAGA---AAAGTAAATGTAAACGGCAAAAGAGAAAATATTA 7150
Db 2912 AAAAGAAAGATATAGATGTATTAAGTGAAGAGTAAAGAGAGTGCCTACACTTTTA 2971
Qy 7151 ACCAAAACAGCAAAATGCAAGAAATCGAGAGCAGCAATCGAGCCAAATGCTTGTGA 7210
Db 2972 TGAAGCTGTGAACGCGAGAAAGAAAGGCGCAATCAATTAACGAAATTTTGAA 3031
Qy 7211 ATTAATTTGGAACGCTGTAGAAAGATAGAAAATTTCTGAGAAAAGCAAGATTT 7270
Db 3032 ATTGAAGAAATGCAAGTGAAGAAATGTAAGAAATGTGCAAGAAATTAAGAAATTA 3091
Qy 7271 TAAAACTTGAAGCAGATTAACAAGAAACAAGTAAATAAGTAATGATCTACGAAA 7330
Db 3092 ACGAAATGTATTTAATTACTGTATTAGATTAAGTAAAGTAAAGCAATTAAGCGGAG 3151
Qy 7331 AATCTTAACATCAGCAGGATTTCTACAGAAATCTTCTGTAAGCGGATAGAGAG 7390
Db 3152 AAAGTTTGAAGAAATGAAGTGAATGAATTTTGTAGTGAATATTTGATATAGTGA 3211
Qy 7391 ATACTAGGAGAAAGAAATTAAGCCATTGTGAAGCTTGTATATTAATGGAAAAATG 7450
Db 3212 AAGGAATACAAAGAAATTTATTAACAGGTATG-----TTTCGAAGTATGAACAACGTA 3265
Qy 7451 TAGATTTTACAAAGAGGAAAGAAATTAATCACTTCTAGTGGTGTGGAACTGAG 7510
Db 3266 TAGTAATCCAAATGAGAAAGAAAGGTGATTTGAATGAAATGTGGTTAGTTCAATTTAG 3325
Qy 7511 GTCTGCTTCGCAATCAGAAACAGTGGCAGTTACAAATATTAAGAAATTTCCGAGTTA 7570
Db 3326 ATTAATATAGAAATATAGAAAGAGTTTATTAATATATAGAAAATATTTCAAGTACG 3385
Qy 7571 CTGTGAAAATTTCTTTTGAAGCAGCTGAAAAGTAAATGTAGATCGGATAT----- 7625
Db 3386 AAGGTGTTCAAGAAACTGTAACTGAAATGTAGAAACAAATGTATATGTGATGTGATG 3445
Qy 7626 ----TTCAGGAATATGTTGCTTTTAAACAGCATATCAAGGCTCTGTAAGAGCAATGGGAAATAG 7681
Db 3446 TTCTGTATAGAAAGATCAATTTTGAAGATATTAATTAAGGACGAGGGGTTGAAGAA 3505

Qy 7682 GAGCTCCTATGCAATTAATTTCTAATGGAAGATCAAAATATCAGTATTAATAATTTCTA 7741
Db 3506 TGTTTTAAATTTGAAAGATGTATTTAAAGTAAAGTGAATTAATCTGTAGAAAGAA 3565
Qy 7742 AGCTATTAAGAAAAATATTTGATGTATTTGTAAAGATTAATTCGAATTTGAGAGCGAAG 7801
Db 3566 TTAAGGATGAACCGGTTCAAAAAGAGGTAGAAAAGAACTGTAGTATTTATGAAGAA 3625
Qy 7802 CAAAAGATTAAACGTAGAGCGGTAGCTCCGAGCAATTAATCTCAAAAGCAAGAAATG 7861
Db 3626 TCGAAGAAATATTTGTAGATGTATTAAGAGAAAGAAAGAAATTTAAACAGCAAGTGA 3685
Qy 7862 AAATGAATTCAGAGTTGAAATTTGAGAGAGTATTTCAATGAAGAAATAGATTACTA 7921
Db 3686 TAGATGCAATGAAGAAATCCATGAATATCTTCAGATTTCTAAAGAAAGAACTGAATCTA 3745
Qy 7922 GCCCTTTAAAGAAATTTGAAAGAAATCAATGTCAAGTGGMAAAGAAACAGAGTGA 7981
Db 3746 TTAAGATTAAGAAAGAAAGATTTTCACTAGTGTGMAAGAAAGTTCAAGCATATATG 3805
Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGCATAGCAGGGCAG 8027
Db 3806 ATGAAGTGTGAGAAAGTTTGAATGAATGAATAATGAGAGAGGATTAATGAAGATG 3865
Qy 8028 AATTAATTCAGAAAGAAAGATCCGGAAGCTCTTAATTTGAAGTATGACAAATCCGG 8087
Db 3866 CTGTTGAATTAATATGACATTTACTAGCAAACTTAATGAAGAACTCAAGGCTTTAAATGAAG 3925
Qy 8088 AAGAATATTTTTCATGCAAGATATGTGAATTTGGAAGCAACATTAATGAAGTAAAC 8147
Db 3926 TGAAGGCAATTTAATTAAGATATGGAAGAAATTAAGAAATTAAGAAAGC--ATTATC 3983
Qy 8148 ACAGTTTCTAAAGAGATTAACGTTCTGTATTTGGAAGAGTTGGAATCAACAGGCAGA 8207
Db 3984 AGAAGTTCTAAAGAAATTAATGATGCAAAAGATATCATTAAGAAAGTATTTGMAA 4043
Qy 8208 ACTACTGCTGAGGTTAAATCTATGTAAGTGAAGAGAAATTTGTCAGAACAA 8267
Db 4044 GGAACATGATTAACGACGATTTGATGMAAGTTTGAAGATTAAGATGTGACAGAA 4103
Qy 8268 TCGATTAATGCAATTTCTAAAGTAAAGGTTTGAAGTAAAGTAACTGTAAATC 8327
Db 4104 CAAGATCGAAAAGTATCTGATTTAAAGATCTTGAAGAGATATTTAAAG----- 4156
Qy 8328 TTCTGTATATCAGAAATGAGAGAAATGTCGGAAGCAGAGTGAATCTTCAACG 8387
Db 4157 --AAGTAAAGAAATCAAGAACTTGAAGTGAATTTTAAGAAATTAATTAAGAAATTA 4214
Qy 8388 ACAAGTATTAATGAAATCCGTAAGTCTGTTTACGAAGCAAGATTATGAAGATTA 8447
Db 4215 AACTATTAAGAAAGATTAATTTGAAGAGAAAGAAAGAAATGAAGAAATCAATTTTGA 4274
Qy 8448 CCAAAAAATATATTTCAAGTCAATGCTCTGCTTTAAATGATACAAAGATAGAG 8507
Db 4275 ATTCGAAGAGAGCTGAAGAAATTAAGATCTTGAAGCATATTTAAAGAAATATC 4334
Qy 4335 TTCATTAGAAAGTTGAAGAAAGAAAGAAATTAAGAAAGTACAGATTAAGAAAGAGT 4394
Db 8568 GAGATCAACAAAGTTAATCTTCAACATGTAATGAGAGAAAGTATCTCACTTGTGTC 8627
Qy 4395 AGAACTAATTAATGAGTGAATGCGATTAAGAAAGGTTTGAAGAAAGATTAAGAAAG 4454
Db 8628 AAAGCTTTGGCTAAATGAAGAAATTAATGAAATGTAAAGAACTGAGAGAGCTTAGT 8687
Qy 4455 AGTATGTA--TTTAAAGAGAGTATTTAGACATGTTAAAGGAGATATGGA--ATTAG 4511
Db 8688 CCGAGCGAAACAGACCGGTTGAATTAACAAGAGTATCAACAGAGCAATGCTTGC 8747
Qy 4512 GATATGATTAAGAAAGTATTAAGATGTAAACAAACAACTTGAGAAAGAGTGAATC 4571

Db 2132 TAGCTCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTAAGCTCCAAATG 2191
 Qy 6380 TAGAGCTTCTTCTGATACCAATATTATPAAAAAG---AAATACCAAGCAAGATTGGA 6436
 Db 2192 TTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGAAGAAATGTTGAAGAAATG 2251
 Qy 6437 AAACACAAATGCTGATGAAGGTTTCGGAAGAAAGCTGAAATTACAGACATCTTAAGC 6496
 Db 2252 TAGCTGAATAATGTTGAAGAAAGTGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGAATAATG 2311
 Qy 6497 AAGCAATTTCCCTCTTTGGAAGTCCGAGTCCGACAGCCGGGTAGAGCCGAGAGTGGCAG 6556
 Db 2312 TTGAAGAAAGTGTAGTGAATAATGTTGAAGAAATGTTAGCTCCAACTGTGAAGAAATCG 2371
 Qy 6557 GAAACGTTTCCGTAATCAATTTGCAAGAAAGCAAGAGTAGATGTGAAGAAAGCA 6616
 Db 2372 TAGCTCAACTGTTGAAGAAATGTTGAGCTCCAAAGTGTGAAGAAAGTGTGCTCCAAAGT 2431
 Qy 6617 TTTTGTAAAAAAAGCTGAGATTACAGCAAAAGTTATAGTCTGTGCAATTGGAATG 6676
 Db 2432 TTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAATAATGTTGAAGAAAGT 2491
 Qy 6677 CCGAGCTCGAGAGTGTGCAAAAAGAGCTGAATTGAGACAGCAGTGGCAATTACCAAG 6736
 Db 2492 TAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAATAATGTTGAAGAAAGTGTAGCTCCAACTG 2551
 Qy 6737 ATGAATCAAAACAGAGACAGAGAGTGAATAATTTCTPAAAT--ATGACTCGAAACAGT 6793
 Db 2552 TTGAAGAAATGTTAGCTCCAAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGT 2611
 Qy 6794 TAGATGTAATAGCAAAAAATGATPAAATCAGTACTGGATCCGTTCAAGCCGAGAGT 6853
 Db 2612 TTGCTGAAGAAAGCTTGCAAAATTTATCGACAACTTTTAAAGTAATTTATTAAGTGGTA 2671
 Qy 6854 GAATCTTGACAGCCGAGATCTGAGTGTCTGCAATATATTGCAATTAAGTGA 6913
 Db 2672 TCGAAACTGAGAAATTAAGAGACAGTATTTAAATGAGATAGAAAGTAAGAAAGAAATG 2731
 Qy 6914 AAACAGATATCATCTAGTACTTTACCTTTCTATCTGATGTAATGTTAAAGCTCTTA 6973
 Db 2732 TAGTCCACCAATATCTAGAAAGCTGTAAGAAAGCTACAGCTGAAGTGTAACTACTTTTA 2791
 Qy 6974 ATPAAATTTGCAATCTTGAAGACGCGGTGAGAGAGCCGAGTCTTGACAGAGTTAACG 7033
 Db 2792 GTTAACTATTAAGAGATCAAGAAATACTATTACTAATGACTATAGAGAAAT 2851
 Qy 7034 GAGTGTCTCTGTAACTATPAAATGTTCTGTAGTCTGAGTTCAATTAATCTG 7093
 Db 2852 TAGAAGAACTCCAGAAATGTTAAGTGTGCTTTAAGAAATACCCAAAGTGAAGAG 2911
 Qy 7094 ATTGCACTTCCGTACAGAG---AAAGTAATGTAAACGCAAAAGAGAAATATATTA 7150
 Db 2912 AAAAGAAAGATTAATGATGTATTTGAAGAGTAAAGAGAGTGTGCTACACTTTAA 2971
 Qy 7151 AGCAAAACAGCAAAATGCAAGATCGAGAGACAGCAATCGAGCCAATGTCTTGTA 7210
 Db 2972 TAGAAACTGTGAAACAGCAGAAAGAAAGCGCAATATCAATTACGGAATATTGGA 3031
 Qy 7211 ATTAATTTGGAACAGCTGTAGAGATAGAAAAATTTCTGAAGAAAGAAAGCAAGAGTT 7270
 Db 3032 ATTTGAAGAAATGCAAGTGAAGATTAAGAAATGTTGCAAGAAATTTAGAGAAATTA 3091
 Qy 7271 TAAAACTTTAGACGAAGTTAAACAAAGAACAGATPAAAAAGTAAATGATCTACGAAA 7330
 Db 3092 AGCAAACTGTATTTAATCTGTATTAATTAAGTAGAGAAACAGTAGAAATTAAGCGAG 3151
 Qy 7331 AAATCTTACATACAGAGGATTTCTACAGAAATCTTCTPAAAAAGCGATAGAGAG 7390
 Db 3152 AAAGTTTGAAGAAATGAATGATTAAGCAATTTTGTAGAAATTTTGAATTAAGTGA 3211
 Qy 7391 ATACTCAGGAGAAAGAAATTAAGCAATTTGTGAAGACTTCTGATATTATGAAAAAATG 7450
 Db 3212 AAGGAATACAAAGAAATTTTATTAACAGGTATG-----TTTGAAGTATGAAGAACAGTA 3265

Qy 7451 TAGATATTACACAGAGACAGAAATATATCACTTCTACTGTGTGTTGGAACTGAG 7510
 Db 3266 TAGTAATCAATCGAAGAAAGGTTGATTTGAATGAATAATGTGTAGTTCGATTTAG 3325
 Qy 7511 GTCTTGCTTCGCATCAGAAACAGTGGCATTACAAATTTAAAGAAATTCGAGTTA 7570
 Db 3326 ATTAATATGAATAATATGAAGAAAGGTTTATTAATTAATTAAGAAATTTCAAGTACTG 3385
 Qy 7571 CTGTGAAATTTCTTTGTGAAGACAGTGAATAAGTAAAGTTAGATCCGAT- 7625
 Db 3386 AAGGTTCAGAAACGTAACTGAACATGTAGAACAAATGTATGTGATGTTGATG 3445
 Qy 7626 ----TACAGAAATGTTGCTTTAAACAGCATATCAAGTCTGTAGAGACTGGAAATG 7681
 Db 3446 TTCTGCTATAGAAATCAATTTTGAATAATTTAAATGAGCAGAGAGGTTGAAGAA 3505
 Qy 7682 GAGTGTCTATGCAAAATTAATTTCTAAATGGAAGTCAAAATATCAGTATTAATAATCTA 7741
 Db 3506 TGTTTTTAATTTGGAAGATGATTTAAAGTGAAGTGTATTTACTGTAGAGAA 3565
 Qy 7742 AGCTATTAGAAAAATTTGTATGTTATTTGAAGATTAATCGAATTTGAGCGGAG 7801
 Db 3566 TTAAGATGAACCGGTTCAAAAAGAGTGAAGAAAGAACTGTAGTATTATGAAGAA 3625
 Qy 7802 CAAAAGATTAAACGTAGAGCGGTAGCGCCGAGCCATTATCTCAAAAGCAAGAAATG 7861
 Db 3626 TGAAGAAATTTGTATGATGTTATTAAGAGAAAGAAAGAAATTTAAACAGCAAGATGA 3685
 Qy 7862 AAATGAATTCAGAGGTTGAAATTTGAAGAGTATTTCAATGMAAGAAATAGAGTAACTA 7921
 Db 3686 TAGATGCAATAGAAATCCATAGAAATATCTTCAGATTTCTPAAAGAGAACTGAATCTA 3745
 Qy 7922 GCCCTTTPAAGGAATTTGAAGAGAAATCATGTCAAAGTGAAGAAAGCAAGTGA 7981
 Db 3746 TTAAGATTAAGAAAGAAAGTGTTCATAGTGTGAAGAAAGTTCAAGCAATGATAG 3805
 Qy 7982 CTGCGAATCTCAAGAGCT-----TCTGTAAGAGCAGTGAAGGAGCAG 8027
 Db 3806 ATGAAGATGTTGAGAAAGTTTGAATTTGAATAATTTGAAGAGAGATTAATGAAGATG 3865
 Qy 8028 AATTAATTCGAAAGCAAAAGATCCGGAAGCTTTATTTGAAAGTTAGTACAAATCCG 8087
 Db 3866 CTGTTGAATAATATGACATTTATACGAAACTATTTGAAGAACTCAAGGTTAAATGAAG 3925
 Qy 8088 AAGAGTATTTTCAATGCAATATGTGAATATGAGCAACATPAAATTAAGATTAAC 8147
 Db 3926 TAGAAGCAGATTTAATTAAGATATGAAAGAAATTAAGAAATTAAGAAAGC--ATTATC 3983
 Qy 8148 AGCAATTTTCAAGCAATACAGTCTGTATTTGGAGAGAGTTGAGTCAACCAAGCAG 8207
 Db 3984 AGAAGATTTCAAGAAATTAATGATGCAAAAGATGATACATTAGAAAAAGTTTGAAG 4043
 Qy 8208 AGCTAGCTGCAAGTAAACTATGAGTGAAGTGTGATGGAAGTGAAGTAAAGTAACTGCTAATC 8267
 Db 4044 GGAACATGATTAAGACGACGTTGATGAAGTTGTAAGTTTAAAGATGTGCAAGAG 4103
 Qy 8268 TTTCTGATATCAGAAATGAGAGAGAAATTCGCGAGACAGAGTGAATCTTCAAGC 8327
 Db 4104 CAAGATCGAAAAAGTATCTGATTTAAAGATCTTGAAGAGATATTTTAAAG----- 4156
 Qy 8328 TTCTGATATCAGAAATGAGAGAGAAATTCGCGAGACAGAGTGAATCTTCAAGC 8387
 Db 4157 --AAGTAAAGAAATCAAGAACTTGAAGTGAATTTTGAAGATTTAAGAAATTTAA 4214
 Qy 8388 ACAAAGTATTAATCTGAATCCGTAGTTGTTTGAAGAAAGCAAGATTAAGAAATTAATGATTA 8447
 Db 4215 AACTATTGAAGAGATTTTGAAGAGAAAGAAAGAAATTAAGAAAGATCAATTTGAAA 4274
 Qy 8448 CACAAAAAATATATTTCAGAAATCAATGCTCTGTTTAAATGATACAAAGATGAAGC 8507
 Db 4275 ATTCGAAGAGAGAGTGAAGAAATTAAGATCTTGAAGCAGATATATTTAAAGAAAGTATC 4334

Db	1977	ATGTTGAAGAAAGTGTACT-----GAAAATGTGAAGAAACGTAGCTCCAACTGTTG	2030
OY	6140	AAAAAGATTGTGTACCAATAGTTCACTCACTAATCTTAAATCTTTATGGCAATG	6199
Db	2031	AAGAAATCTGTAAGCTCCAACTGTTGAAGAAATTTAGTCTCCAACTGTTGTA--GAAGTGT	2087
OY	6200	CCGCTGGTTCCAGACAAGCCGAGTGGCAGAACTGTAAATATCAACMAGTTTATGAG	6259
Db	2088	TGGCTCCAAAGTGTGAAGAAAGTGTAGAAGAAATGTTGAAGAAAGTGTGAAGAAATG	2147
OY	6260	AAACAGAAAGCTCTGTGAGAAATTTCTATATTAATGCAAAACATTATCTGTAAATCAG	6319
Db	2148	TTGAAGAAAGTGTAGCTGAGAAATTTGAGAAAGTGTAGTGAAGAAAGTGTGAAGAAATG	2207
OY	6320	GAGATTACAGAAATTCATCGGATAGTAGGTTCTGTGGTGTGGTGAAGAAATGTAGAG	6379
Db	2208	TAGCTCCAACTGTTGAAGAAATCTTAGCTCCAACTGTTGAAGAAATTTAGTCTCCAACTG	2267
OY	6380	TAGAGAGCTTCTCTGATACCAATATTATTAAG--AAATACCAAGACAAGTGTGA	6436
Db	2268	TTGTAGAAAGTGTGGCTCCAACTGTGAAGAAAGTGTGAAGAAAGTGTGAAGAAAGTGTG	2327
OY	6437	AAACTACAACTCTGATGAAAGTGTGAGAGAAAGTGAATTTACAGCAATTTAAGC	6496
Db	2328	TAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAGAAATG	2387
OY	6497	AAGGAATTTCTCTTTTGGAGTCCGAGTCCGAGAGCCGGGTATGAGCCGGAGTGGCAG	6556
Db	2388	TTGAAGAAAGTGTAGCTGAGAAATTTGAAGAAATCTAGCTCCAACTGTTGAAGAAATG	2447
OY	6557	GAAACGTTTCCGTAAATCAATTTTCAGAGAAAGACGAAGTAGATGTGAAGAAAGCAAGA	6616
Db	2448	TAGCTCCAACTGTTGAAGAAATTTAGTCTCCAACTGTTGAAGAAAGTGTGGCTCCAACTG	2507
OY	6617	TTTTGGTAAAAAAGCTGAGATTACAGCAAAACGTATAGTTCTGTTCATTTGGAATG	6676
Db	2508	TTGAAGAAAGTGTGAAGAAATGTGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTG	2567
OY	6677	CCGCACTCGAGTGGCTGCAAAAGAGCTGGAATTTGAGACGACAGTGGCAATTACCAAG	6736
Db	2568	TAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTCCAACTG	2627
OY	6737	ATGATCAACACGAGACGAGAGTGAAGAAATTTCTAAAT--ATGACTGAAACAAGT	6793
Db	2628	TTGAGAAATTTGTAGCTCCAACTGTTGAAGAAAGTGTAGCTCCAACTGTTGAAGAAAGTGTG	2687
OY	6794	TAGATGTAAATAGCAGAAATATGATTAATTAATCAGTACTGAATCGTTCAAGCCGAGCTG	6853
Db	2688	TTGCTGAAGAAAGTGTGAAGAAATTTATCAGCAATCTTTAATGTAATTTATTAAGTGTGA	2747
OY	6854	GAAATTTCTGACGCCGAGATATCTGGAGTGTCTCTGTCATATATATTTGCCAAATTAAGTAG	6913
Db	2748	TCGAAGTGTGAAGAAATTAAGACAGTATATTAATTAATGAGTGAAGAAATTAAGAAATG	2807
OY	6914	AAACGATATGATATGATCTTACACTCTTCTACTGATGTAAATGTAAAGCTCTTAA	6973
Db	2808	TAGTCAACCAATATTAGAAACGTAAGAAAGTCACTACGCTGAAGTGTAACTCTTTTA	2867
OY	7034	GAGTGGTTTCTGTAACTATTAATAGTTCTGTGATAGCTCGAGTTCACTAATACCTGTG	7093
Db	2928	TGAGAGAACTCCACCAAAATGTATTTAGTGTGCGCTTTGAAATTAACCAAGTGAAGAG	2987
OY	7094	ATTGACTCTCGTAGAG--AAATAGTAATGTAAAGGCAAAAGAGAAAAATATATA	7150
Db	2988	AAAGAAAGAAATTAATGATGTAAATTTGAAGTGAAGAAAGAGTGTGCTCACTTTAA	3047
OY	7151	ACCAAAACAGCAATATGAGGAATTCGAGAGAGACGATCGAGCCAAATGCTTGTGTAA	7210

Db	3048	TAGAAACTGTGGAACAGCGAGAGAAAAGCCGAATACAAATTACGGAATAATTGGAAA	3107
Qy	7211	ATTAATTTTGGAAACGCTGTAGAAAGATAGAAAAATTTCTGAAGGAAAAGAACAGAACTTT	7270
Db	3108	ATTTAGAAGAAAATGCAAGTAGAAAGATGATGAAAAATGTTCGACGAATTTAGAAAAATTAA	3167
Qy	7271	TAAAACTTTAGACGAAGTTAACAAAGAACAGTTAAAAAGTTAAATGATGCTACGAAAA	7330
Db	3168	ACGAAACGTATTTTAAATACGTATTTAAATAAAGTAGAGGAAAACGTAGAAAATTGCGGAG	3227
Qy	7331	AAATCTTACATCAGCAGGTATTTCTACAGAAATACCTTCTGTAAAAAGCGATAGAGAG	7390
Db	3228	AAAGTTTAGAAAAACAAATGAATGATTAAGACTTTTTTACTGAAAATTTTGATTAATGTA	3287
Qy	7391	ATACTCAGGAGAAAGAAATTAAGCCATGTGAAGACTCTGATATTATTGAAAAAATG	7450
Db	3288	AAGGAATACAGAAAATTTATTAAACAGTATG-----TTTCGAAGATAGAAAACAGTA	3341
Qy	7451	TAGATATTACACAGAGACACAAATATATCACTTCTACTCGTGCTTTGGAACTCCAG	7510
Db	3342	TAGTAATCCATCAGAGAAAAGGTTGATTTGAATGAAAATGTGGTAGTTGCGATTTTAC	3401
Qy	7511	GTCTTCCTCCGACATCAGGAACAGTGGCAGTTACAAATATTAAAAAATTCGCGATTA	7570
Db	3402	ATTAATATAGAAAATATGAAAAGAGTTTATTAAATTAATAGAAAATTTTCAAGTACTG	3461
Qy	7571	CTGTGAAAAATCTTTTGTGTGAAGCAGCTGAAAAAGTAAATGTTAGATCGGATAT----	7625
Db	3462	AAAGTGTTCCAGAAAACGTGTATCTGAACATGTAGAACAAAATGTAATGTGATGTTGATG	3521
Qy	7626	---TACAGAAATGTTGCTTTAAACAGATATCAAGGTCTGTAGAGCATTTGGGAATAG	7681
Db	3522	TTCCGTATAGAAAAGATCAATTTTTCAGAAATTTAAATGAGCAGAGGCGTTGAAAGAA	3581
Qy	7682	GAGCTGCGCTATAGCAGAATTAATCTATATGGAAGTCAAATATCAGATTTAAAAATCTA	7741
Db	3582	TGTTTTTAAATTTGGAAAGATGATATTTAAAGTGAAAGATGATATTTCTGTAGAGAAA	3641
Qy	7742	AGCTATTAAGCAAAAAATATATGATGTTATTGTAAAAAGATTAATTCGAAATTGAGCCGAG	7801
Db	3642	TTAAGAGATGAACCGGTTCAAAAAAGAGTAGAAAAAGAAACCTGTAATTTATTGAAAGAA	3701
Qy	7802	CAAAAAGATTAACCGTAGACCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAAGATG	7861
Db	3702	TGGAAGAAAATATGTATGATGATATTAGAGGAAAAAAGAAATTTAACACACAAAGTA	3761
Qy	7862	AAATGAATTCAGAGTTGAAATTTGAGAGATATTTTCAATGAGAGAAATAGAGTAATCTA	7921
Db	3762	TGATGACAGTAGMAAATCCATAGMAAATCTTCAGATTTCTTAAGAGAAACCTGATATTA	3821
Qy	7922	GCCCTTCTAAAGGAATTTGAAAGAAATCAATGTCTCAAGTGAAAAAAGAAAAACAGACTGA	7981
Db	3822	TTAAAGATTAAGAAAAAAGATGTTTCACTAGTTGTGGAAGAGTTCAAGACATGATATGG	3881
Qy	7982	CTGTGAATCTCAAGACT-----TGTAGAGCAGTAGCAGGGCAGG	8027
Db	3882	ATGAAGATGTGGAAGAAATTTAGAAATGAAAAATATGAAAGAGAGCTTAATGAAGATG	3941
Qy	8028	AATTATTTCCGAAGCAAAAGATCCGGAAGCTCTTATTGAAAGTTACTACAAAATCCGG	8087
Db	3942	CTGTGAAAATTAATATCAATTAACCAACTTATTGAAAGAACTCAAGAGTTAAATGAG	4001
Qy	8088	AAGAAGTATTTTTCATGCGATATATGTAATATGGAACCAACATCAATTAATGAAGTAAAC	8144
Db	4002	TGAAAGCAGATTTAAATTAAGATATGAAAAATTTAAAGAAATTTGAAAAAAGC--ATTATC	4055
Qy	8148	ACAGATTTCTAAAGCAGTAACAGTTCTGTATTGGAGAGAGTTGAGTCAACCAAGCAGA	8207
Db	4060	AGAAAGTTCTAAAGAAATTAATGATGCAAAAGATATCAATTAGAAAAAGTTATTGAAGA	4119
Qy	8208	AGCTACTGTCGACAGTAAACATATGTATGAAAGTTGAGAGAGAAATTTGTTCAAGACAA	8266
Db	4120	GGAAACATGATATTAACGACGCTTGGAATGAAGTTGTAAATTTAAAGATGTGGAAGAA	4179

8268 TCGATTGATGCAATTTCTAAAGTAGAAAGTTGGATGAGATTAAGTAATGCTAAATC 8327
4180 CAAGATCGAAAAAGATCTGATTTAAAGATCTTGAGAGAGATATATTTAAAG----- 4232
8328 TTCTGTAGTATCAGAGAAATGAGAGAAATTCGCCGAGCAGAGATGTAATCTTCTACAGC 8387
4233 --AGTAAAGAAATCAAGAACTTGAAAGTGAATTTTGAAGATTTTAAAGAAATTA 4290
8388 ACAAGGTAATCTGAATCCGTAGTTCGTTAGCAAGCAAGTTTATGAAATTAATGATTA 8447
4291 AACCTTTGAAACAGATATTTTGAAGAGAAAAAGAAATGAAAGATCATTTTGA 4350
8448 CACAAAAAATATATTTTCAAGATCAATCTCTCTTAAATGATATACAAAGATGAGC 8507
4351 ATTCAGAGAGAGAGTGAAGAAATTAAGATCTTGAGCAGATATTTTAAAGAGATTC 4410
8508 GAATATGAAATCTTTAGCCGTAGCCGCTGTCATGCACAGAGAACAAAGCATTTAC 8567
4411 TTCAATTAAGAAAGTGAAGAAAGAAATTTGAAGAGATACGAAATTTAAAGAGAGT 4470
8568 GAGATCAACAGATTAATCTTCAACATGTAATGAGAGAAAGTATCTCACTTCGTGC 8627
4471 AGAATATTAATTAAGTGTGATGCGCATATTAAGGTTTGAAGAGATGATTTGAAGA 4530
8628 AAAAGCTTTGCTAAAAATGAAATTTATGAAATGTAAAGAGACCTGAGAGCCTTAGT 8687
4531 AGTATAGTGA-TTTAAAGAGAGATATTTAGACATGTAAAGGAGATATGGA--ATTAG 4587
8688 CGAGCCGGAACAGCAGCCGTTGAAATTTATCAAGAGTACTACAGAGCATTTGCTGC 8747
4588 GATATGATTAAGAAAGTTGAAGATGTAAACAAACTTGAGAGAAAGATGGAATC 4647
8748 AGGAAATTTGGGAAATTTGAGATTAATTAAGAAAGTGAAGAGATTAATACATTTGAG 8807
4648 CTTAAAGATGTTTATCTAGTGCATTTAGCATGATGAGAGAAAGAAAGAAAGAAAG 4707
8808 AGTCAACGAGACGAGAACCAAGAGAGTCTTGCAGAAAGATGATTTCTGTA 8867
4708 AAAAGCTCAAAACCTTAAGTTGAGAGATATTTAAAGAGAGGTTTAAAGAGAAC 4767
8868 TACAATTTCAAGGAGAAACAAATTCATCCATGAGATTAAGCCAGAAATTTGGAACCG 8927
4768 AAAGAAAAAATTAACAAAAAAGAAAGTGAAGTTGATTAATTAAGATTAAGAGAAC 4827
8928 AAGTGAATTTAGATGCTTTGAATGAATTTGATTAATTAATTAATTAATTAATTAAT 8987
4828 TGAATATGAGAGTTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 4887
8988 CTATGCTGAATTTGATATTTGAAATTTGATTAATTAATTAATTAATTAATTAATTAAT 9047
4888 AGAAGAGATTAAGAGAGATTAAGTTGAGATTAATTAATTAATTAATTAATTAATTAAT 4947
9048 AGCCAAATTCGAGAGACATGCTATTTGAGAACTGAGAAAAAGAAATTAATTAATTAAT 9107
4948 AGGTGAAGCAAGATGAAGTTATGATTTAATAGTCCAAAAAGAGAAAGCATTTGAAA 5007
9108 TACAAGAGCAAAAGTAAATTTCTTGAAAAAGAGACGCTGACGCTGCAATATC 9167
5008 GGTAAAGGAAAAAGAAAAATTTGAAAAAGAGGTTGAGGTTGCTTAA 5067
9168 GAATGTACATTTCCATGAGATGATATTAATTAATTTGCAAAAGCATATGATCTTC 9227
5068 AAAACCGTAGAGAGATGATATTTCTTCAAAAAATTTGATTAAGAGAGTTGATTAAGA 5127
9228 TCAATTAATTAACCAAAATTTCAAAATTAATTAATTAATTTAGTCATCA 9274
5128 AGTATCTAAAGCTTGAATCAAAAAATGATGTTACTATATGTTTAA 5174

RESULT 24
AR131271
LOCUS AR131271 6152 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 1 from patent US 6191270.
ACCESSION AR131271
VERSION AR131271.1 GI:14119596
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
AUTHORS Unknown.
TITLE 1 (bases 1 to 6152)
JOURNAL Malarial pre-erythrocytic stage polypeptide molecules
PUBMED 6191270-A 1 20-FEB-2001;
FEATURES
source location/Qualifiers
1..6152
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 1.2%; Score 116.2; DB 6; Length 6152;
Best Local Similarity 41.7%; Pred. No. 4.6e-06;
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

5126 GTTCTGATGAACCGAAGCTTTAGTATGATTTCTGAGTTTGAAGAGTAACTTTCA 5185
1017 GTTTGAAGAAAGTATAGCTTCAAGTTGATGAAGAGTATGATTCAGATTGAAGAA 1076
5186 ATGTAGATGCAAAAGATCAAAACAAATTAATTAACAA--TTGCCGAAATGCAATGAG 5242
1077 ATGTAGCTCAACGTGTTGAAGAAATCGTAGCTCCAGAGTTGTAAAGAGTGGCTCCA 1136
5243 GAAAGCGCTGAGAGTTGAGCAACAGTTGCTCATACAAATTTGAAAAACATCAGTTA 5302
1137 GTGTTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGAAGAAATGTTGAAGAA 1196
5303 TAGCTATTTGTAACAAAGTAAATTAACAGCGGAAATGATCAAGATGAAAAATATCA 5362
1197 GTTAGCTGAAAGTGTGAAGAAAGTGAAGTGAAGAAAGTGAAGAAAGTGAAGTGA 1256
5363 ATGTAGCTCAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5422
1257 ATGTGAAGAAATGTAGACTCCAACTGTTGAAGAAATGTAGCTCCAACTGTTGAAG 1316
5423 AAGAGCCTCTGTCAAGAGACTTCTGCAAGTACTTCAATTAAGACAGTTTCTTCTC 5482
1317 TTTAGCTCCAAAGTGTGTAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGAAGAA 1376
5483 ATGTGATCAACATGATTTGACAAAGATTTTGAAGAGAAATTAATTAATTAATTAAT 5542
1377 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAA 1436
5543 AGCCAAATGTTAATGTTTGTAGCTGAAGAAATGAGCTCAAGTGTGCAAAATGCGAC 5602
1437 GTGTAGCTGAAGATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAG 1496
5603 TTTCCGAGCAAGTGAAGCAAGCTGAGAGAGCTGAGAGAGTGAATTAATTAATTAAT 5662
1497 ATGTGAAGAAATGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAACTGTTGAAGAA 1556
5663 AAAATACTTGTGCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5722
1557 TTTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAG 1616
5723 GC---AAATTCATTCATCTAATTAACAAATTTGAAATTTGAGAGTGTGAGCTGAG 5779
1617 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGA 1676
5780 GAGCTGAGTGAAGAGTCTGTAGCAGTGAATTAAGATTTGAATTAATTAATTAATTAAT 5839
1677 GTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAG 1736
5840 TAAATCATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5899
1737 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGA 1796

Qy	5900	TAAATGCTAAATTATGACGAAACAGTGTCTGGAGTGGCCCGTCACAGCAATAGAGCTCGAA	5959
Db	1797	TCGTAGCTCCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTGTAGCTCCAA	1856
Qy	5960	CCAGTGTGAATGAAGAAATTACAGAGATCTACAAAGATATGTATAAAGATTTCTACGTGATTTG	6019
Db	1857	GTTGTTGTGAAGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAAGTGTGAAGAA	1916
Qy	6020	CTAAGAAGAAACAGATGATTAATTAATTAATCTCAAGGGCAAGTGAATMAAGTGTGATAT	6079
Db	1917	GTGTAGCTGAAAATGTTGAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGTGA	1976
Qy	6080	AAGTATTCAAAATCTTAATATTAACGAGAATTATCACAAAAAGAAAAATAGTATAT	6139
Db	1977	ATGTTGAAGAAAGTGTAGCT-----GAAAAATGTTGAAGAAATGTGTAGCTCCAACTGTTG	2030
Qy	6140	AAAAAGATTTGTTATCCAAATAGTTCAAGTACTCACTACTTTTAAATCTTTATTTGGCAATG	6199
Db	2031	AAGAAATCGTAGCTCCAACTGTATTAGAGAAATGTGTAGCTCCAAAGTGTGTA--GAAAGTG	2087
Qy	6200	CCGCTGGTTGAGACAAGCCGAGTGGCAGGAACTGTAAATATACAAAGATTTATGAG	6259
Db	2088	TGGCTCCAAAGTGTGAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTAGCTGA	2147
Qy	6260	AAACAGAGCTCTGTAGAAAAATCTATATTTAAATGCAAAACATTATCTGTAAATCAG	6319
Db	2148	TTGAAGAAAGTGTAGCTGAAAATGTGTGAAGAAAGTGTAGCTGAATATGTTGAAGAAATCG	2207
Qy	6320	GAGATTACAGAAATTTCAATCGAGTAGTAGGTTCTGTGGTGTGGTGTGAATGTACAG	6379
Db	2208	TAGCTCCCAACTGTGAAGAAATCGTAGCTCCAACTGTGTGAAGAAATGTGTAGCTCCAAAGTG	2267
Qy	6380	TAGAGCTTCTTCTGATACCAATATTAATAAG--AAATACCAAGCAAGATTTGAA	6436
Db	2268	TTGTAGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTG	2327
Qy	6437	AAACTACAAATGTCTGATGAGGTTTCGAGAGAAAGCTGAATTTACAGCAGATTTCTAAGC	6496
Db	2328	TAGCTGAAAATGTTGAAGAAAGTGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGA	2387
Qy	6497	AAAGCAATTTCTCTTTTGAAGTCGAGTCGACAGCCGGGTACGAGCCGAGATGGCAG	6556
Db	2388	TTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAATCGTAGCTCCAACTGTGTGAAGAAATCG	2447
Qy	6557	GAAACGTTTCCGTAATCAATTTTGGACGAAAGACGAGTAGTAGTGGAAAGAACAGAA	6616
Db	2448	TAGCTCCCAACTGTGTGAAGAAATGTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTG	2507
Qy	6617	TTTTGGTAAAAAAGCTGAGATTACAGCAAAACGTTATAGTTCTGTTCGCAATTTGAAATG	6676
Db	2508	TTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTAGCTGAATAATGTTGAAGAAAGTG	2567
Qy	6677	CCGCACTGGAGTGGCTCCAAAAGAGCTGGAATTTGACAGACAGTGGCAGTTACCAAG	6736
Db	2568	TAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAACTG	2627
Qy	6737	ATGAATCAAAACAGAGACGAAGTGAAGAAATTTCTAAATTT--ATGACTGGAAGCAAGT	6793
Db	2628	TTGAAGAAATGTGTAGCTCCAACTGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGTG	2687
Qy	6794	TAGATGTATATAGCAAAAATGAGATTAATAATCAGTACTGGAATCGGTTCAAGCCGAGCTG	6853
Db	2688	TTGCTGAAAACGTTGCAACAAATTTATCAGCAATCTTTTAAAGTAAATTTATTAAGTGTGA	2747
Qy	6854	GAAATTTGACAGCCGAGATTTGGAATGTGTTCTGTCAATATATTTGCCAATAAAGTGA	6913
Db	2748	TCGAATCTGAGAGAAATTAAGACAGATATTTAAAGAAATGAGAAAGTGAAGAAAGTGA	2807
Qy	6914	AAACAGATTCGATATAGTACTTTTAACTCTTCTACTGATGTAAATGTAAAGTCTTTA	6973
Db	2808	TAGTCAACCAATACTAGAAAAAGTGAAGAAATCTACAGCTGAAGTGTAACTACTTTTA	2866
Qy	6974	ATTAATTTTGAATTCCTTTGACGCGGTGAGAGCCGACAGTCTTTCACAGCATTTACG	7033

Db	2668	GTAACTATTAGAGGAGATCCAGAAATACTATTACTAAATGATACATATGAGAAAAAT	2927
Oy	7034	GAGTGCTTTCTGTAACTATATAATAGTTCTGTGATAGCTCGAGTTCACAATACTCTG	7093
Db	2928	TAGAAAGAACTCCAGAAAATGTATTAAAGTCCGCTTTAGAAAAATACCAAAGTAAGG	2987
Oy	7094	ATTTGACTCCGTACGAGA--AAAGTAAATGTAAACGCGAAAGAGAAAAAATTTA	7150
Db	2988	AAAAGAAAGAAATGATGTATGTATTGAAAGTAAAGAAAGAGTGCTACCCCTTAA	3047
Oy	7151	AGCAACAGCAGAAATGCGGAATCGGAGAGCAGCAATCGAGCCAAATGCTCTGTAA	7210
Db	3048	TAGAACTGTGGAACAGCGCAGAAAGAGCCGAAATCAATTTACGGAATATTGAA	3107
Oy	7211	ATTAATTTTGGCAACAGCTGTAGAAAGATGAAAAAATTCGAAAGAAAGAACAGAGTTT	7270
Db	3108	ATTTAGAAGAAAAATGCAAGTAGAAAGTAAATGAAATGTTGCAGAAATTTAGAAAAATTA	3167
Oy	7271	TAAAACTTTAGACGAAGTTAACAAAGAACAGATAAAAAGTAAATGATGCTACGAAA	7330
Db	3168	ACGAACCTGTATTAAATACGTATTAGATAAAGTAGAGAAACGTAGAAATTTGCGGAG	3227
Oy	7331	AAATCTTAACATCAGCAGTATTTCTACAGAAATCTCTGTAAAGCGGATGAGAGG	7390
Db	3228	AAAGTTTAGAAACATAGAAATGGATATTTTAGGAAATATTGATTAATGTA	3287
Oy	7391	ATACTCAGGGAGAGAAATTAACCCATTTGGAAGACTTCTGATATTTATGGAAAAATG	7450
Db	3288	AAGGAATACAGAAAAATTTATTAAACAGTATG-----TTTCGAAGTATAGAACAGTA	3344
Oy	7451	TAGATATTACAACAGAGCACAAGATATATACCTTCTACTGTGGTGTGGGAACTGCAG	7510
Db	3342	TAGTAATCAATCAGAAAGAAAGGTTGATTTGAATGAAAAATGTGGTTAGTTGCATTTTAC	3401
Oy	7511	GCTTGCTCCGCATCAGGAACAGTGCAGTTACAAATATTAAGAAATTCGCGAGTA	7570
Db	3402	ATAATATAGAAAAATATGAAAGAGTTTATTAAATTAATGAAAAATTTTCAAGTACTG	3461
Oy	7571	CTGTTGAAAAATTTCTTTGTGAAACAGCTGAAAAAGTAAATGTTAGATCCGATAT----	7626
Db	3462	AAGGTGTTCAAGAACTGTAACTGAACATGTAAGAACAAATGTAATGTGATGTGATG	3521
Oy	7626	----TACAGAAATGTTGCTTTAAACAGATATCAAGTCTCTGAGAGCATTTGGAAATAG	7681
Db	3522	TTCCGCTATGAAAGATCAATTTTATAGGAATATTAATGAGGCGAGGTTGAAAAA	3581
Oy	7682	GAGCTGCTATGCAGAATTTAAATCTAATGGAAGTCAAAATATCAGTATTTAAAAATCTA	7741
Db	3582	TGTTTTTAAATTTGGAAATGTATTTAAAGTGAAGTGAATTACTGTAGAGAA	3641
Oy	7742	AGCTATTAGAAAAAATATGTAGTATTTGTAAAGATPAAATCGAAATTGAGCGGAG	7801
Db	3642	TTAAGGATGAACCGGTTCAAAAAGAGTGAAGAAAGAACTGTTATTTATGAGAA	3701
Oy	7802	CAAAAGATTAAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAGAAATG	7861
Db	3702	TGGAAAGAAATTTGTATGATGTAATAGAGGAAGAAAAAGAAATTTACAGACAAGATGA	3761
Oy	7862	AAATGAATTCAGAGTTGAAATTTGAGAAGATATTTCAATGAAGAAAAATAGAGTACTA	7922
Db	3762	TAGATGCAGTAGAAATCCATAGAAATATCTTCAGATTTCAAGAAACAACTGAATCTA	3821
Oy	7922	GCCCTTAAAGGAATTTGGAAGAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGTA	7981
Db	3822	TTAAAGATTAAGAAAAAATGTCTTCAATGTTGTTGAAGAAATTCACACATGATATAG	3881
Oy	7982	CTGCTGAATCTCAAGAGCT-----TCTTAGAGCAGTAGACAGGGCAGG	8021
Db	3882	ATGAAGATGTGAAGAAATTTTAGAATGAAAAATATGAAGAGAGTTTAAATGAAGATG	3941
Oy	8028	AATTATTTCCAGAGAAAGATGCCGGAAGCTCTTATTGAAGTATATACAAATCCGG	8081

Db 3942 CTGTTGAAATTAATGACATTACTACGAACTTAATTGAAAGAACTCAAGATTAAATGAAG 4001
 Qy 8088 AAGAGTATTTTTCATGACAGATATATGATATGAGACACACATTAATGAATGAATAC 8147
 Db 4002 TAGAAGCAGATTTTAAATTAAGATATGAGAAATTTAAAGAAATTAGAAAAAGC-ATTATC 4059
 Qy 8148 AGCAGTTTCTAAAGCAGTACAGGTTCTGTATTGGAGAGATTGGAGTCCAGGCGAGA 8207
 Db 4060 AGAAGATTCTAAAGAAATTAATGATGCAAAAGATATACATTAGAAAAAGTTATTGAAAGA 4119
 Qy 8208 AGCTACTGTCGAGGTAAATCTATGATGAGTATGAGAGAAATTTGTCAGAAACAA 8267
 Db 4120 GGAACATGATATACGACGACCTTGATGAGATTGATTAATTAAGATGTCGAAAGAGA 4179
 Qy 8268 TCGATTGATGCAATTTCTAAAGTAGAAGTTGGATGAGATGAAGTAAAGTACTGCTAAATC 8327
 Db 4180 CAAGATCGAAAAAGTATCTGATTTAAAGATCTTGAAGAGATATATTAAAG----- 4232
 Qy 8328 TTCTGTAGTATCAGAAATGAGAGAAATTCGCGAGCAGAGTGAATCTTCTACAGC 8387
 Db 4233 --AAGTAAAGAAATCAAAAGAACTTGAAAGTAAATTTTGAAGATTTAAAGAAATTAA 4290
 Qy 8388 ACAAGTATATGATCCGTAGTTCGTTTACGAAAGCAAGTTATGAAAAATATGATTA 8447
 Db 4291 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATGAAAAAGATCATTTTGAAGA 4350
 Qy 8448 CACAAAAAATATATTTTCAGAGTCAATGCTCTGCTTAAATGATACAAAGATGAAGC 8507
 Db 4351 ATTGCAAGAGAGAGTGAAGAAATTAAGATCTTGAAGAGATATATTAAAGAAAGTATC 4410
 Qy 8508 GAATATGAAATCTTTAGCCGTAGCCGCTGTGCATGCACAGGACAAACAAAGCATTTAC 8567
 Db 4411 TTCATTAGAGTGAAGAGAAAAAAATTAAGAGATGACGATTTAAAGAAAGAGT 4470
 Qy 8568 GAGATCAACAAAGTTAACTTCAACACTGTAATGAGAGAAAGTATCTCACTTCGTCG 8627
 Db 4471 AGAACATTAATTAAGTGTGATGCGCATTAAGAGTTTGAAGAGATGATTGAAAGA 4530
 Qy 8628 AAAAGCTTTGCTAAAAATGAATAATTAATGAATGTAAGAGAGTGAAGAGCTTACT 8687
 Db 4531 AGTATATGA- TTTAAAGAGAGTATATTAAGATGTTAAAGAGAGATATGGA-ATTAG 4587
 Qy 8688 CGAGCGGAGACAGAGCCGTTGAAATTAATCAAGAGTACTACAGAGCATTTGTCG 8747
 Db 4588 CGATATGATTAAGAGAAAGTTGAGATGTAACAACAACTTGAGAGAAAGATTGATC 4647
 Qy 8748 AGGAATTTGGGAAATTTGAGATTAATTAAGAAAGATGCAAGAGTATATCATTTGTA 8807
 Db 4648 CTTAAAGATGTTTATCTAGTCAATTAGCATGATGAGAGAAACAAATGAAACAAAGAA 4707
 Qy 8808 AGTCAACGAGACGAGAACCAAGAGAGTCTGTGCGAAAGATGGTATTTCTGTAAAAA 8867
 Db 4708 AAAAGCTCAAAAGCTTAAGTGGAGAGATATTTAAAGAGAGGTTAAAGAAAGAAC 4767
 Qy 8868 TACAAATTCAGGGGAGAAACAAATTCATCATTGAAGATTAAGCAGAAATTTGTGAACCG 8927
 Db 4768 AAAAGAAAAAATTAACAAAAAAGTAAGTTGATTAATTAAGATTAAGAGAACCAAGA 4827
 Qy 8928 AAGTAAATGATGATGCTTGAATGAATTTGATGATTAATTAAGATTAAGAGAAAGTGTG 8987
 Db 4828 TGAATATGATGAGAGTGAATTAAGATGAAGATTAAGAGAGATGTAAGAGAGATAT 4887
 Qy 8988 CTATGTGGAATTTGATATGGAATGTTGATGATTAATTAAGATTAAGAGAAAGTGTGA 9047
 Db 4888 AGAAGAGATATGAAGAGATTAAGATTAAGATTAAGATTAAGATTAAGAGATAT 4947
 Qy 9048 AGCCAAATTCGAGAGACATGATTTGTAGAAACTAGTGAACAAAGATATCAAGCAT 9107
 Db 4948 AGGTGAAGCAAGATGAAGTATAGATTTAATAGTCAAAAAAGAGAAACGATTTGAAAA 5007
 Qy 9108 TACAGAGCAAAAGTAAATATTTCTTGAAAGAGAGAGCTGCAGCTGCAGCTCAATATC 9167
 Db 5008 GGTAAAGGAGAAAAAGAAAAATTAAGAAAAAAAGTTGAAGAGGTGTTAGTGTCTTAA 5067

Qy 9168 GAATGTACACATTTCCATGATGATGATATTTAAAAATTTGGCAAGAGATGATCTTC 9227
 Db 5068 AAAACAGCTAGACGAGATTAAGATATGTTCAAAAAATTAATGAAGAGTTGATTAAGA 5127
 Qy 9228 TCAATTATTAACCAAAAAATTCAAAAAATTAATTAATTAATTAATTAATTAATTA 9274
 Db 5128 AGTATCTTAAGCTTTAGATCAAAAAAGATGTTACTAATGTTTAA 5174

RESULT 25
 AC164627
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS

AC164627 192275 bp DNA linear ROD 27-JUL-2005
 Mus musculus 10 BAC RP23-382A18 (Rosewell Park Cancer Institute
 (C57BL/6J Female) Mouse BAC library) complete sequence.
 AC164627
 AC164627.4 GI:71274204
 HTG.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Mus.
 1 (bases 1 to 192275)
 Muzny,D., Adams,C., Agbai II,O., Allen,C., Albrooks,S., Archer,P.,
 Arredondo,H., Bandaranaike,D., Bangura,L., Beltran,B., Beltran,R.,
 Berarducci,A., Biswal,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
 Cadore,J., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,
 Chacko,J., Chahour,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
 Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,
 Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De
 Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H.,
 Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durkin,K.,
 Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
 Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
 Fowler,G., Fu,Q., Fun,B., Garcia,A., Garcia,R., Garner,T.,
 Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D.,
 Gonzalez-Garay,M., Guevara,M., Holder,M., Haaland,M., Heberlein,K.,
 Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Hawlak,P.,
 Hawes,A., Hawking,E., Hayes,S., Hemphill,L., Hernandez,J.,
 Hines,S., Hitchens,M., Hodgson,A., Hognes,M., Holline,B.,
 Howell,L.T., Huylk,S., Hume,J., Imo,K., Jackson,A., Jackson,L.,
 Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalatus,K., Kelly,S.,
 Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,F.,
 Leal,S., Lee,K., Lee,S., Legall,F.I., Lemon,S., Lewis,L., Li,B.,
 Li,Y., Li,Z., Linell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyange,D.,
 London,P., Lopez,J., Lorenzowa,L., Lorado,R., Luk,T., Mach,R.,
 Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,B.,
 McElliand,H., McPherson,D., Mercado,C., Metzker,M.,
 Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Mundasa,M.,
 Murray,D., Nazareth,L., Ngo,D., Nguyen,N., Norwig-Bastan,B.,
 Nott,A., Nwaleleleh,O., Obregon,M., Ochli-Okorle,C., Odeh,E.,
 Okunolu,G., Okunolu,K., Parker,D., Pasternak,S., Patel,B.,
 Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Phan,T.,
 Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quiriz,J.,
 Rabata,D., Rachlin,B., Reigh,R., Ren,Y., Reuter,A., Richards,S.,
 Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W.,
 Santibanez,J., Santos,R., Savery,G., Scherer,S., Shen,H., Shen,Y.,
 Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R.,
 Svalter,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R.,
 Trejos,Z., Umami,K., Varjo,C., Verduzco,D., Villaseana,D., Virk,D.,
 Volkov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J.,
 Wei,X., Wheeler,D., Williams,G., Williams,R., Morley,K., Wright,R.,
 Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L.,
 Zhang,Z., Zhou,D., Weinstock,G. and Gibbs,R.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE

2 (bases 1 to 192275)
 Morley,K.C.
 Direct Submission
 Submitted (25-JUN-2005) Human Genome Sequencing Center, Baylor
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 192275)

AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (25-JUL-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4 (bases 1 to 192275)
AUTHORS	Worley, K.C.
JOURNAL	Direct Submission
COMMENT	Submitted (27-JUL-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA On Jul 27, 2005 this sequence version replaced gi:71069494. Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.
	The repeat regions shown were identified using RepeatMasker by Adrian Smit.
	Sequence similarities were identified using Powerblast by Jinghui Zhang.
	Exon/intron boundaries of identified genes were chosen if there were canonical splice junctions across the splice junctions.
	----- Genome Center ----- Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Drafting Center Code: BCM Contact: hgsc-help@bcm.tmc.edu, location/Qualifiers 1. 192275 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="10" /clone="RP23-382A18" 1..1246 /rpt_family="L1Md_F3" 1254..2558 /rpc_family="L1_Mm" 2647..2775 /rpc_family="L1MA4" complement(2801..3079) /rpc_family="L1Md_F3" 3080..3106 /rpc_family="L1_Mus4" 3107..3287 /rpc_family="B2_Mm1" 3288..3454 /rpc_family="L1_Mus4" 3466..3896 /rpc_family="L1_Mus4" complement(4207..4333) /rpc_family="B1_Mur1" complement(4475..4800) /rpc_family="L1Mc4" 4801..4903 /rpc_family="B1_Mur3" 4936..5045 /rpc_family="PB1D9" 5046..5082 /rpc_family="(GN)" complement(5256..7066) /rpc_family="L1Md_F3" 7219..7337 /rpc_family="MT2A" 7402..7707 /rpc_family="MT2B2" 8118..8178 /rpc_family="ID4" 8368..8473

```

repeat_region /rpt_family="B1_MuB1" 8474. .8502
/rpt_family=" (CAAAC) n"
repeat_region /rpt_family="B1_MuB1" 8503. .8523
/rpt_family="B1_MuB1" 8526. .8581
/rpt_family=" (GAAA) n"
repeat_region /rpt_family="B1_MuB1" 9764. .9791
/rpt_family="AT_rich" 10466. .11393
/rpt_family=" (CAG) n" 11412. .11678
/rpt_family=" (CAG) n" 11700. .11882
/rpt_family=" (CAG) n" 11895. .12077
/rpt_family=" (CAG) n" 12090. .12272
/rpt_family=" (CAG) n" 12285. .12466
/rpt_family=" (CAG) n" 12479. .12661
/rpt_family=" (CAG) n" 12689. .12873
/rpt_family=" (CAG) n" 12885. .13063
/rpt_family=" (CAG) n" 13106. .13291
/rpt_family=" (CAG) n" 13319. .13503
/rpt_family=" (CAG) n" 13539. .14083
/rpt_family=" (CAG) n" 14549. .16252
/rpt_family=" "MuBHAL1" 16253. .16751
/rpt_family=" "MuBHAL1" 17002. .17122
/rpt_family=" "B1_MuB1" 17705. .17829
/rpt_family=" "B1_MuB2" 17945. .18040
/rpt_family=" "RSINEL1" 19007. .19222
/rpt_family=" (CAG) n" 19451. .19510
/rpt_family=" "CT-rich" 19679. .19763
/rpt_family=" (TA) n" 21667. .21708
/rpt_family=" (TCTG) n" 21924. .21950
/rpt_family=" "AT_rich" 23702. .24290
/rpt_family=" "11M4_P2" 26473. .26801
/rpt_family=" "Lx" 27692. .27728
/rpt_family=" "CT-rich" 27729. .27767
/rpt_family=" (TTTG) n" 27768. .27893

```

Db 11197 AGAGCAGAGCAGGAGCAGCAGCAGCAACAGAGCAGGAGCAGCAGCAACAG 11256
Qy 5670 TTCTGCACATATATAAATAAGTACTCAATATGTACGAATGCTTTGGTAAAAAGCAATC 5729
Db 11257 AAGACGAGAGAGAGCAGCAGCAGCAGCAACAGAGCAGCAGCAGCAACAGAGCAGC 11316
Qy 5730 TCATTCTATTTAAACAAATTGGAATTGGAGCTGGAGTTGGAGCTGGAGAGCTGGAGT 5789
Db 11317 AGAGCAGCAGAGTGTAGCAACAGAGCAGCAGTAAACAGAGAGCAGCAGAGCAGCTGC 11376
Qy 5790 GACAGGTTCTGTAGCAGTGAATTAAGATTTTAATATATACGATAGCAGATTAATCATGC 5849
Db 11377 AGCAGAGAGCAGCAGGAGGAACAGTGGTAAACACAGCAGCAGCAACCAAGAGCAGC 11436
Qy 5850 AAAAAATCACTGCGAAGGGAATATGTCCGAGTTATTACAGAGTCTGATGCCGATTTGCTAA 5909
Db 11437 AACAGAGAGCAGCAGCAGCAACAGAGCAGCAGCAGTAAACAGAGCAGCAGAGCAGC 11496
Qy 5910 TTATGCAAGAACAGTGTCTGGAGTGGCCCGTGCAGCAATAGAGCCTCAACCAAGTGTGA 5969
Db 11497 AGTAGTAGCAACAGAGCAGCAGCAACAGAGCAGCAGAGCAGCAGCTCCAGCAGCAGC 11556
Qy 5970 TGAATATTAAGAGATCTACAAAAGCATATGTAAAGATTTACAGATGATTGCTAAAGAGA 6029
Db 11557 AGCAACAGCAGCAGGAGCAGCAGCAACAGAGAACAGCAGCAGCAGCAGAGCAACGT 11616
Qy 6030 AACAGATGATTAATTACTACTCAAGGCAAGTAGATTAAGTGTATGATTAAGTATTCA 6089
Db 11617 GGTAAACAAGCAGCAGCAGCAACAGACAGCAGCAGAGAGCAGCAGCAACAGAGCAGC 11676
Qy 6090 AAATCTTAATTAATTAACGAAGACTTATCACAAAAAGAAAAATAGTAAATAAAAAGATT 6149
Db 11677 AGGTCCAGCAGGTCCAGCAGGTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCA - 11734
Qy 6150 TGTTCACATATGTTGCTGCTACTCATCTTTAAATTTTATTGGCAATGCCGCTGGTTC 6209
Db 11735 -GAAGCAACAGAGAGCAGGAGGAACAGTGGTAAACAGCAGCAGCAGCAACAGAGC 11793
Qy 6210 AGGACAGCCGAGAGTGGCAGGAAGTGTATATCAACAAGTTTATGAGAAACAGAGC 6269
Db 11794 AGCAGCAGGAGCAGCAGGAGGAACAGTGGTAAACAGCAGCAGCAGCAACAGAGCAGC 11853
Qy 6270 TCTTGTAGAAAACTTATTAATGAATGCAAAACATTAATCTGTAAATCAGAGATTAAC 6329
Db 11854 AGCAGAGCAGCAGCAACAGAGCAGCAGCAGTCCAGCAGGTCCAGCAGCAGCAGCAACAG 11913
Qy 6330 GAATTTAAATCCGAGTGTAGTGTCTGTGTGTGTGTGTGAATGTAGAGTGTAGAGCTTC 6389
Db 11914 AGCAGAGCAACAGTAGAAGCAACAGGGGAGCAGAGGAACAGTGTAAACAGAGCAGC 11973
Qy 6390 TTCTGATACCAATATTAATAAAGAAATACCAAGACAAAGATTGAAAAATACATGTTC 6449
Db 11974 AGCAGACACAGCAGCAGCAGCAGCAGCAGAGAGAAAGTGTGTAAACAGCAGCAGC 12033
Qy 6450 TGATGAAGTTTCGAGAAAGAGCTGAAATTACAGCAGATTTCAAGCAAGAAATTTCTTC 6509
Db 12034 AGCAACCAAGAGCAGCAGCAGCAGCAGCAACAGAGCAGCAGGTCCAGCAGGTCCAGC 12093
Qy 6510 TTTTGAAGTCCGAGTGCAGCAGCCCGGGTGAAGC-----CGAGTGGCAGAAACCT 6563
Db 12094 AGCAGAGCAACAGGAGCAGCAGCAACAGAAAGCAAGGGGGGCGCAGGAGGAACGT 12153
Qy 6564 TTCCTGAATCAATTTGCAAGAAAGCAGAAAGTAGATGTGGAAGAGCAAAATTTTGGT 6623
Db 12154 GGTAAACAAGCAGCAGCAGCAACAGACAGCAGAGAGAGCAGCAGGAGAACTGGT 12213
Qy 6624 AAAAAAGCTGAGATTACAGCAAAACGTTATAGTTCTGTGCAATTGGAATGCGCAGT 6683
Db 12214 AACAAACAGCAGCAGCAGCAACAGAGCAGAGAGAGCAGCAGCAACAGAGCAGCAG 12273
Qy 6684 CGAGTGTGCTGCAAAAGAGCTGGAATTGAGCAGCAGTGGCAGTTTACAAAGATGAATC 6743
Db 12274 TCACAGAGGTCCAGCAGCAGCAGCAACAGAGCAGCAGCAACAGAAACCAACAGGGGA 12333

Qy 6744 AACACGAGCAGAGTGAATAATTTCTAAATTAATGACTCGAAAACAGTTAGATGTAAAT 6803
Db 12334 GCAGAG- GAACAGTGTAAACAAGCAGCAGCAGCAACAGAGCAGCAGAGCAGCAGC 12392
Qy 6804 AGCAGAAAAATAGATTAATTAATCAGTACTGGAATCGTTCCAGCCGAGCTGGAATTTTGC 6863
Db 12393 AGGAGGAACAGTGGTAAACAAGCAGCAGCAGCAACAGAGCAGCAGCAGCAGCAGC 12452
Qy 6884 AGCCGAGTATCTGGAGTGTCTGTCAATTAATTTGCAATAATAGTAAGAAACGATAT 6923
Db 12453 AACAGAGCAGCAGTCCAGAGGTCCAGCAGCAGCAGCAACAGAGCAGCAGCAACAGC 12512
Qy 6924 CGATCATAGTACTTTTACACTCTTCTACTGATGTAATGTAAAGCTCTTAATTAATTTTC 6983
Db 12513 AGAAGCAACAGAGCAGCAGAGGAACAGTGTGTAACAACAGCAGCAGCAACAGAGC 12572
Qy 6984 GAATTCCTTGAACAGCCGTTGAGAGCCGAGCTTGGACAGCATTAACCGAGTGGTTTC 7043
Db 12573 GGCAGCAGAGCAGCAGCAACAGAGCAGCAGGTCCAGCAGCAGCAGCAGCAAC 12632
Qy 7044 TGTTAACCTAATAATAGTTCTGTGATGCTCGAGTTCAATTACTGATTTGACTTC 7103
Db 12633 AGAGCAGCAGCAGCAGCAAGCAACAGAGAGCAGAGGAACAGTGTAAACAACAGC 12692
Qy 7104 CGTACGAGAAAAAGTAAATGTAAACGCAAAAAGAGAAAAATTAAGCAACAGCAGC 7163
Db 12693 AGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 12752
Qy 7164 AAATGCAAGAAATCCGAGAGCAGCAATGAGCCATGTCTTGGTAAATTTTGGAAAC 7223
Db 12753 AGCAGAGAGCAGCAGCATAGCAGCAGCAGCAGCAGCAAAAATATCAGCAGAGC 12812
Qy 7224 AGCTGTAAGATGAATAAATTTCTGAGAGAAAGGAACAGAGTTTAAAAAATTTAGA 7283
Db 12813 AATGAACACAGCAGAGGAAGCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAACAGC 12872
Qy 7284 CGAAGTTAAACAAGAAACAAGATTAATAAAGATTAATGATCTGAAAAAATCTTACAATC 7343
Db 12873 AACGCAACCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 12932
Qy 7344 AGCAGTATTTCTACAGAAATCTTCTGTAAAGCGGATAGAGAGATCTCAGGAGGA 7403
Db 12933 AGCAGCAGCATAGCAGCAGTAGCAAAAATATCAGCAACCAACAGAGCAGCAGCAGGA 12992
Qy 7404 AGGAATTAAGCCATTTGGAAGACTTGTGATTAATTTGGAATAATGTAGATTAACAAC 7463
Db 12993 AGAGCAGCAGAGCAGCAGCAGAGCAACAGAGCAACAGAGCAGCAGCAGCAGCAAC 13052
Qy 7464 AGAGACAAAGATATATCACTTCACTGTGTGTGTGTGTGGAACAGAGTCTTGTCCGC 7523
Db 13053 CAGAGCAGCAAGAAAGCAGAAATCAACAGAGAGCAGCAGAGAGAGTGTAAACAGCAGC 13112
Qy 7524 ATCAGAAACAGTGGCAGTTACAAATTAATAAGAAATTCGAGATTAAGTTGAAAAATTC 7583
Db 13113 AGCAGCAGCAGCAACCAAGCAGCAGAGAGCAGCAGCAGCAACAGCAACAGCAGC 13172
Qy 7584 TTTTGTGAACAGCTGAATAAGTAAATGTAGATCGGATTTACAGGAATGTTTGCCTT 7643
Db 13173 AGCAGCAGCAGCAGTGTGAGCAGGAAATGACAGCAGCAGCAGCAGAGGAACAGAGC - 13229
Qy 7644 AACAGCATATCAAGGTCTGTAGAGCAATTGGGAATAGAGCTGCTTATGCAATTTAA 7703
Db 13229 AGCAGAGCAGCAGTCCAGCAGCAGCAGCAACAGAGCAGCAGCAACAGAGCAAC 13289
Qy 7704 TTTTAATGAAGATCAATATCAGTATTAATAATTTCTAAGCTATTAGAAAAAATTTGA 7763
Db 13280 AGAAGAGCAGAGAGGAACAGTGTAAACAACAGCAGCAGCAGCAACAGAGCAGCAGGA 13349
Qy 7764 TGTATTGTAAGATTAATCGGAATTGAGCCGAACCAAAAGATTAACTGCTGTAGAGC 7823
Db 13350 AGCAGCAGCAAGAGCAGCAGAGGAAGCAGCAGAGAGCAGCAGAGAGCACTTAGAGAGCAGC 13409

QY	7824	GGTAGCTGCCGGAGCCATTATCTCAAGGAAGAAATGAAATGAATTCAGAGTTGGAAT	7883
Db	13410	AGGTGACGACGAAACGAGCAGCAGCAGCATGACGACGAGGACGAGCAGCAACGACAA	13469
QY	7884	TGAGAAAGATATTTTTCATGTAMAAAAATAGATTAAGTACTAGCCCTTCTTAAAGGATTTGGAAG	7943
Db	13470	CTAGACGACGACAGGTGTACAGCAGGTGTGACAGCAACGAGGACGACGACGAGGACGACAC	13529
QY	7944	AGAAATCATATGTCAAAGTGTAAAAAGAAAAAGAGTGACTGCTGAATTCAGAGAGCTTC	8003
Db	13530	AGCAACCTGTGACGACGACGACGAGCAACAGCAGCATGACGACGACGACGACAAATATCAAC	13589
QY	8004	TGTAGGAGCGATGACGAGGGCAGGAATTTATTCGAAAGCAAAAAGATGCCGGAAGCTTTA	8063
Db	13590	AGGAGCAACAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	13649
QY	8064	TTTGAAGATTAAGTACAAAAATCCGGAAGAAAGTATTTTTCATGACGATATGTGATATGGA	8123
Db	13650	AGCAGCAACAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAG	13709
QY	8124	AGCAACCATPAAAAATGAAAGTAAACAGCATTTCTAAAACGTAAACAGGTTCTGTATTGGG	8183
Db	13710	AGCAGCAGGACGACGTGTGTAACACAGCAG-----CAGGAGCAACAGGAAAAACGACAG	13763
QY	8184	AGGAGTTGAGATCCACCAAGCAGAGATCTGTGTCGACGAGTAAAACTATGTTGAAGTTGA	8243
Db	13764	AGCAGCAACATCAACTGTAGAGCAGAGAGGTGTCTGACGACCAACGAGCAGCAGCAGC	13822
QY	8244	GGAAGAAATTTGTTCAAGAACAAATCGATTGAATGCAATTTCTTAAAGTAGAAGTTTGA	8303
Db	13824	AGCAGGAGCAGCAGAACCAACAGGACGACGACGACGACGACGACGACGACCAACGCACTTG	13883
QY	8304	TGAAGATPAAAGTAACTGCTAAATCTTTCTGTATGTAATTCAGGAAATGAGAGAGGAAATTCGCG	8363
Db	13884	AGCAGCAGCAGCAGCAGCAACACAGAGCAGCAGCAGGACGACGACGACGACCAACGAGACAGC	13943
QY	8364	AGCAGAGGATGAATCTTCTACAGCAAA	8392
Db	13944	AGGACGACGACGACGACGACGACCAAAAA	13972

AC164629/C	RESULT 26				
LOCUS	AC164629/C				
DEFINITION	AC164629	203395 bp	DNA	linear	HTG 17-JUL-2005
ACCESSION	AC164629				
VERSION	AC164629.3	GI:70912408			
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FUL1TOP; HTGS_ACTIVEPIN.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;				
	Sciuognathi; Muridae; Murinae; Mus.				
	1 (bases 1 to 203395)				
REFERENCE	Muzny D., Adams C., Agbali II.O., Allen C., Alsbrooks S., Archer P.,				
AUTHORS	Arrondondo H., Bandaranaike D.B., Bangura L., Beltran B., Beltran R.,				
	Berducci A., Blawie K., Blyth P., Bonham L., Buhay C., Burch P.,				
	Chadree I., Canada A., Cardenas V., Carter K., Cavazos I.,				
	Chadree J., Chhrotr M., Chavez D., Chen A., Chen G., Chen R.,				
	Cheng M.-T., Chu Y., Clerc K., Cockrell R., Coyle M., Cree A.,				
	Curry S., Dai W., Davila M.L., Davis C., Day-Carroll L., De				
	Ande C., Delgado O., Denson S., Deramo C., Ding Y., Dinh H.,				
	Donlin J., McCauley S., Dugan-Rocha S., Dunn A., Durbin K.,				
	Dezinda D., Egan A., Escotte M., Espinosa V., Eugene C., Fa M.,				
	Fernandez S., Fernando P., Flagg N., Forbes L., Foster B.,				
	Forster G., Fu Q., Fuh E., Garcia A., Garcia R., Garner T.,				
	Gastin C., Gensch S., Ghose S., Gill R., Gonzalez D.,				
	Gonzalez-Garay M., Guevara W., Holder M., Haland W., Heberlein K.,				
	Hall B., Hamid H., Hamilton K., Harbes B., Harris R., Haviak P.,				
	Hawes A., Hawking E., Hayes S., Hemphill L., Hernandez J.,				
	Hines S., Hitchens M., Hodgson A., Hogues M., Hollins B.,				
	Howell L.T., Huylk S., Hume J., Imo K., Jackson A., Jackson L.,				

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Jacob,T., Jiang,H., Johnson,B., Johnson,R., Kalatsky,K., Kelly,S.,
Kevs,I., Khan,Z., King,L., Kovac,C., Kowis,A., Kowitz,C., Lara,F.,
Leal,S., Lee,K., Lee,S., Legall,F.I., Lemon,S., Lewis,L., Li,B.,
Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Lyranage,D.,
London,P., Lopez,J., Lorenzini,E., Lozano,R., Luk,T., Madu,R.,
McNishwari,M., Mahoney,C., Mallory,K., Mansouri,D., Martinez,E.,
Mchaeland,H., McPherson,J., Mercadado,C., Metker,M.,
Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Munidasa,M.,
Morray,D., Nazarith,L., Ngo,D., Nguyen,N., Norwis-Eastanga,E.,
Nott,A., Nwokelemeh,O., Obregon,M., Ochi-Okorie,C., Odoh,E.,
Omwunon,G., Okwunon,K., Parker,D., Pasternak,S., Patel,B.,
Patel,V., Paul,H., Perez,A., Perez,L., Petrostino,J., Pham,T.,
Primus,E., Pu,L.-L., Puzao,M., Qin,X., Quinn,A., Quintoz,J.,
Rabate,D., Rachlin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S.,
Rives,C., Rodriguez,F., Rojasy,A., Ruiz,S.J., Sana,M., Sanders,W.,
Sanlibanaz,J., Santos,R., Savery,G., Scherer,S., Shen,H., Shen,Y.,
Slason,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R.,
Swatsek,A., Taylor,E., Taylor,T., Thomas,N., Thorm,R., Thornton,R.,
Tejero,Z., Usmani,K., Varigo,C., Verduzco,D., Villanasa,D., Virk,D.,
Volkov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J.,
Wei,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R.,
Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L.,
Zhang,Z., Zhou,J., Weinstein,G. and Gibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Unpublished

Direct Submission

2 (bases 1 to 203395)

Worley,K.C.

Direct Submission

Submitted (25-JUN-2005) Human Genome Sequencing Center, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 203395)

Worley,K.C.

Direct Submission

Submitted (17-JUL-2005) Human Genome Sequencing Center, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jul 17, 2005 this sequence version replaced gi|68300450.

NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	53145:	contig of 53145 bp in length
53146	53245:	gap of unknown length
53246	159347:	contig of 106102 bp in length
159348	159447:	gap of unknown length
159448	162068:	contig of 2621 bp in length
162069	162168:	gap of unknown length
162169	166874:	contig of 4706 bp in length
166875	166974:	gap of unknown length
166975	170586:	contig of 3612 bp in length

Db	136066	AAACGCAACCTGAGACGACGAGCTGACGACGACACACAAACAGACGAGACGATTGACG	136007
Qy	7344	AGCAGGTATTTCTACAGAAAGTACTTCTGTAAACCGGATGAGGAGTACTCAGGAGAG	7403
Db	136006	AGCAGACCATGAGCAGCAGTGTAGCAAAATATCAGCAAGCAACAGGAGCGACGACGAG	135947
Qy	7404	AGGAATTAAAGCCATTGTGGAAGACTTCTGATATTATTGGAAGAAATGTAGATTACAC	7463
Db	135946	AGAACACACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	135887
Qy	7464	AGAGGACAAGAAATATATCACTTCTACTGTGGTTTGGAACTGACAGGCTTGTCCGC	7523
Db	135886	CAGACGACAGAAACAGCAAGATTAACAAGAGGAGACGACGAGAGCAGTGTAAACGACG	135827
Qy	7524	ATCAGGAACAGTGGCAGTTACAAATATTAAAAAATTCGAGTTACTGTTAAAAATTC	7583
Db	135826	AGCAGCAGCAGCAGCAACAGACGACGAGAGGAGCAGCAGCAGCAACAGCAGCAGC	135767
Qy	7584	TTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGTCGGATATTACAGGAATGTTGCTTT	7643
Db	135766	AGCAGCAGCAGCAGGTTGACGACGAGAAATTAACACAGCAGCAGCAGCAGCAGCAGC	135710
Qy	7644	AACAGCATATCAAGGCTCTGTAGAGCATTTGGGAAATGAGGCTGCTATGAGAAATTA	7703
Db	135709	AGCAGAGACAGCAGGTTCCAGCAGCAGCAGCAACAGAGCAGCAGCAGCAGCAGCAAC	135650
Qy	7704	TTCTAATGGAAGATCAAAATATCATGATTAAAAATTTCTAAGCTATTAGAAAAAATATTGA	7763
Db	135649	AGGAGAGCAGAGGAAACAGTGTAAACAACAGCAGCAGCAACAGCAGCAGCAGCAGCAG	135590
Qy	7764	TGTTATTGTAAAAAGATTAATCCGAAATTGAGACCGAAGCAAAAGATTAAACGTAGAGC	7823
Db	135589	AGCAGCAGAGAAAGACGACGAGAAAGCAGCAGAGCAGCAGAGAGCAACTGAGGAGCAGC	135530
Qy	7824	GGTACTGCGCGAGCCATTATCTCAAAAGCAAAAGATGAATATTCAGAGTTGAAAT	7883
Db	135529	AGGTGACAGACAAACAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	135470
Qy	7884	TGAGAAAGATATTTTCAATGAAGAAATAGATTAATGACTAGCCCTTCTTAAAGAAATTGGAAG	7943
Db	135469	CTGACACAGCAGCAGGTGCGACGAGTGTACACAGCAACAGAGCAGCAGGAGCAGCAAC	135410
Qy	7944	AGAAATCAATGTCAAAGTGAAGAAAGAAACAGAGTGACTGTAATCTCAAGAGCTTC	8003
Db	135409	AGCAACGTGACGACGACGACGACGACCAACAGCATGACGACGACGACGACGACGACGAC	135350
Qy	8004	TGTTAGAGAGATGACGAGGCGCAGGAATTATTTCCAGCAAGAAAGATCCCGGAAGCTCTTA	8063
Db	135349	AGGACCAACAGAGCAGCAGCAGCAGCAACAGCAGGAGCAGCAGCAGCAGCAGCAGCAGC	135290
Qy	8064	TTTGAAGTTAGTACAAATTCGGAAGAGTATTTTTCATGCAATATGTGAATTGGA	8123
Db	135289	AGCAGCAACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	135230
Qy	8124	AGCAACACATAAATGAAAGTAAACAGCAGTTTCTAAAGCAGTAAACAGTTCTGTATTGG	8183
Db	135229	ACACAGCAGCAGCTGTAAACAACAGCAG-----CAGAGCAACAGGAAAAACAACAG	135176
Qy	8184	AGGAGTTGGAGTCAACCAAGCAGAACTACTGTCCGAGTAAATCTATGTATGAATTTGA	8243
Db	135175	ACACAGCAATCAACTTGTAGCAGCAGCAGGTTGTGACGACCAACAGCAGCAGCAGCAGC	135116
Qy	8244	GGAAAGAAATTTGTTCAGAAACAATGATTTGAATGCAATTTCTTAAAGTGAAGTTTGA	8303
Db	135115	AGCAGAGCAGCAGAAAGCAACAGGACGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGC	135056
Qy	8304	TGAAGATAAGTAACTGCTAAATCTTCTGTAGTACAGAAATGAGAGAAATTCGCGG	8363
Db	135055	AGCAGCAGCAGCAGCAACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	134996
Qy	8364	AGCAGAGTGAATATCTTTCTACAGCAAAA	8392
Db	134995	AGGAGCAGCAGCAGCAGCAGCAAAAA	134967

AC153503/

LOCUS AC153503 203787 bp DNA linear HTG 05-JUN-2005

DEFINITION AC153503 clone RP23-107G22, WORKING DRAFT SEQUENCE, 6 unordered pieces.

ACCESSION AC153503 GI:16955894

VERSION AC153503.5

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP; HTGS_ACTIVERPIN.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1. (bases 1 to 203787)

AUTHORS Muzny, D., Adams, C., Abbat, I. O., Allen, C., Alabrooks, S., Archer, P., Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R., Berarducci, A., Biewald, K., Blyth, P., Bonham, H., Buhay, C., Burch, P., Cadoree, J., Canada, A., Cardenas, V., Carter, K., Cavazos, J., Chacko, J., Chadrour, M., Chavez, D., Chen, A., Chen, G., Chen, R., Curry, M., T. Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Croall, L., De Anda, C., Delgado, O., Denison, S., Dexamio, C., Ding, Y., Dinh, H., Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Duphin, K., Duinda, D., Egan, A., Becotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gaskin, C., Gensch, S., Ghouse, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlen, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, A., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B., Howell, L. T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalatus, K., Kelly, S., Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lafer, F., Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y., S., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorensbawa, L., Lozada, R., Luk, T., Madu, R., Maidensthat, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E., McCelland, H., McPherson, J., Mercado, C., Metker, M., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Muidasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastang, E., Nott, A., Nwokilemech, O., Obergson, M., Ochli-Okorie, C., Odeh, E., Okunonu, G., Okunonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Phan, T., Primus, E., Pu, L., Puazo, M., Qin, X., Quinn, A., Quiroz, J., Rabats, D., Rechlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rivas, C., Rodriguez, F., Rojars, A., Ruiz, S. J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savary, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Usmani, K., Varzo, C., Verduzco, D., Villaseana, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, D., Wu, J., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, J., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.

TITLE Unpublished

JOURNAL Direct Submission

REFERENCE 2 (bases 1 to 203787)

AUTHORS Worley, K. C.

TITLE Direct Submission

JOURNAL Submitted (14-DEC-2004) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 203787)

AUTHORS Worley, K. C.

TITLE Direct Submission

JOURNAL Submitted (05-JUN-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jun 5, 2005 this sequence version replaced gi:16955894.

COMMENT ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM


```

QY 6924 CGATCATAGTACTTTACACTCTTCTTACGTGTAAATGTAAAGCTTTATATTAATTTTC 6983
Db 136182 AGAAGCAACAGAGGACGACGAGAGAAACAGTGTAAACACAGACGACGACCAAGAC 136123
QY 6984 GAAATTCCTTTGACGCGGTGAGAGCCGCGACAGTCTTTCGACGACGTTTACCGAGTGTTC 7043
Db 136122 GGCACAGAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 136063
QY 7044 TGTTAACATTAATTAATGTTCTGTGTAGTCTCGAGTTCAATTAATCTGATTTGACTTC 7103
Db 136082 AGGACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 136003
QY 7104 CCTACGAGAAAAGTAAATGTAAACGCAAAAGGAAAATTTAATCAACAAACGACGAC 7163
Db 136002 AGCAGACGACCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135943
QY 7164 AAATCGAGAAATCGAGGACGACGACATCGAGCCAAATGTCTTGTAAATTAATTTTGAAC 7223
Db 135942 AGCAGAGGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135883
QY 7224 ACCTGTAGAGATAGAAAAATTTCTGAAGAAAAAGAACAGAAAGTTTAAAACTTTAGA 7283
Db 135882 AATTAGAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135823
QY 7284 CGAAGTTAACAAAGAACAGATTAATAATGATGCTAGGAAAAAATCTTCAATC 7343
Db 135822 AACAGCACTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135763
QY 7344 AGCAGTATTTCTACAGAGATTAATCTGTAAAAAGCGATTAAGAGATTAATCAAGGAGA 7403
Db 135762 AGCAGACGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135703
QY 7404 AGGAATTAAGCCATTTGTGAAGACTTCTGATATTTATTTGAAAAATGTAGATTAATCAAC 7463
Db 135702 AGAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135643
QY 7464 AGAGACAAAGATTAATATCACTTCTCTGTGTGTTGGGAACCTGCAGGCTTGTCTCCGC 7523
Db 135642 CAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135583
QY 7524 ATCAGAAACAGTGGCGATTAACAATTAATTAATAAATTCGAGTTACTGTGAAAAATTC 7583
Db 135582 AGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135523
QY 7584 TTTTGTGAAGACGCTGAATAATGTTAGATCGATATTATTAACAGAAATGTTGCTT 7643
Db 135522 AGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135466
QY 7644 AACAGCATATCAAGTCTGTAGAGCATTTGGAAATGAGAGCTGCTTATGCAATTTAA 7703
Db 135466 AGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135406
QY 7704 TTCTAATGGAAGTCAATATATCAATCTTAATAAATTTCTAAGCTTTTGAAGAAAATATTTGA 7763
Db 135405 AGGAGAGACGAGGAGAAACAGTGTAAACACAGACGACGACGACGACGACGACGACGAC 135346
QY 7764 TGTATTTGTAAAGATTAATCGGAATTTGAGCGGAGCAAAAGATTAACCTGTAGAGC 7823
Db 135345 AGCAGACGACGAGAAACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135286
QY 7824 GGTAGTGCCTGAGCCATTAATCTCAAAAGCAAAATGAATGAATTTCAAGGTTGAAT 7883
Db 135285 AGGTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135226
QY 7884 TGAGAAAGATTTTCAATGAAGAAAATAGATTAATCTAGCCCTTTTAAAGAAATTTGGAAG 7943
Db 135225 CTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135166
QY 7944 AGAATCATATGTCAAAAGTGAAGAAAAGAAAACAGAGTGAATCTGATCTCAAGAGCTTC 8003
Db 135165 AGCAACGTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135106
QY 8004 TGTAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 8063

```

```

Db 135105 AGGACCAACAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135046
QY 8064 TTTGAAGTTGATCAAAATCCGGAAGATTTTTCATGACGATTAATGTCAATATGGA 8123
Db 135045 AGCAGCAACAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 134986
QY 8124 AGCAACATATAAATGAAGTAAAGTACAGGTTTCTAAGACGATTAACGTTCTGTATTGGG 8183
Db 134985 AGCAGCAGAGCAGTGTAAACAGCAG-----CAGAGCAACAGAGAAACAGCAG 134932
QY 8184 AGGAGTTGAGTCAACCAAGCAGAACTTCTGCTCAGAGTAAACTATGTGAAGTTGA 8243
Db 134931 AGCAGCAACATCAACTGTGACGACGACGACGACGACGACGACGACGACGACGACGAC 134872
QY 8244 GGAAGAAATTTGTTTCAAGCAATCAATGATTAATTTTAAAGTAAAGTTTGA 8303
Db 134871 AGCAGAGCAGCAGAAACGAGACGACGACGACGACGACGACGACGACGACGACGAC 134812
QY 8304 TGAAGATAAGTAACTGCTAAATCTTGTAGTATCAGGAATTTGAGAGAAATTTGCCG 8363
Db 134811 AGCAGCAGCAGCAGGAAACAAACAGACGACGACGACGACGACGACGACGACGACGAC 134752
QY 8364 AGCAGAGTGAATTAATCTTCTACAGCACA 8392
Db 134751 AGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 134723

```

```

RESULT 28
AC141261
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
JOURNAL
TITILE
AUTHORS
JOURNAL
COMMENT

Project Information
Center Project Name: 2732500
Center clone name: RPCI-11_1335J10
-----
Summary Statistics
Consensus quality: 185790 bases at least Q40
Consensus quality: 186288 bases at least Q30
Consensus quality: 186526 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 187243; sum-of-ctdigs estimation
Quality coverage: 12.9 in Q20 bases; agarose-fp estimation
Quality coverage: 12.06 in Q20 bases; sum-of-ctdigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```


Db	35552	AGCTGGCGAGAGGATGACCAATCAGAGTTGCGATGCTGTGGCGAGAGGATGAC	35611
Oy	7539	AGTTACAAATATTTAAAGAAATTCGGAGTTACTGTTGAAAATCTTTTGTGAAGCAGC	7598
Db	35612	AGCATCAGCAGTTGCGATGCTGTGGCAGAGAGATGACGATCAGCAGTTGAAATGAC	35671
Oy	7599	TGAAAAGTAATGTTAGATCGATATTACAGAAATGTTGCTTTAACGCAATACAGG	7658
Db	35672	TGTGGCAGGAGGATGACCAATTCAGTAGCTGTGGCAGGAGTAGCAGCAGCTGT	35731
Oy	7659	TCCGTGAGGAGCATTTGGGATTAGACCTGCTATGCGAAATTAAATTCTAATGGAATC	7718
Db	35732	TGCGTAGCTGTGGCAGAGAGATGACAGCAGCAG--TTGCAGTAGCTGTGGCAGGAG	35790
Oy	7719	AAATATCAGTATTTAAATCTTAAGCTATTAGGAAAAATTTT---GATGTTATTGPA	7774
Db	35791	GAGTAGCAGCATCAGCAGTTGCGATGACTGTGGCAGAGAGATCCAGCATCAGTAGTTG	35850
Oy	7775	AAGATTAATCGGAATTGAGCGGAAACAAAGATTAAACGTAGAGCGGTAGCTGCCG	7834
Db	35851	CAGTAGCTGTGGCAGGAGGATGACAGCAGCAGCATGTCAGTAGAGGTGTGGCAGGAG	35910
Oy	7835	GAGCCATTATCTCAAAAGCAAAGATGAAATGAAATTCAGAGTTGAAATTTGGAAGATA	7894
Db	35911	TAGCAGCAGTTGCGATGCTGTGGCAGGAGATGACACAGTTGCGATGAGCTGTGGCAG	35970
Oy	7895	TTTTCAATGAAGAAATAGAGTAATCTAGCCCTTCTAAAGAAATTGGAAGAAATCAATG	7954
Db	35971	GAGAGTAGCAGCAGTTGCGATGACTGTGGCAGAGAGATGACGATCAGCAGTTGCCAG	36030
Oy	7955	TCAAAGTGAAGAAAGAAACAGAGTGACTGCTGAATCTCAGAGAGCTTCTGTAGAGCAG	8014
Db	36031	CAGCTGTGGCAGGAGGATGACAGCANTCAGCAGTTGCAATGAGTGTGGCAGGAGTAGAGC	36090
Oy	8015	TAGCAGGGGCGAGATTTATTTCCGAGCAAAGATGCCGAGCTTTATTGAAAGTTA	8074
Db	36091	CATCAGCAGTAGCAGTAGCTGTGGCAGAGAGATGACGATCAGCAGTTGCGATGAGCTG	36150
Oy	8075	GTAACAAATCCGGAAGATATTTTTCATGCAATATGTAATGAAAGCAACATTA	8134
Db	36151	TGGCAGGAGGATGACAGCATCAGCACTTCAATGCTGTGGCAGAGAGATGACAGCAT	36210
Oy	8135	AAATGAAGTAACAGCAGTTTCTTAAACAGATTAACAGTTTCTTATTTGGAGAGTTGGAG	8194
Db	36211	CAGCAGTTGCGAGCAGCTGTGGCAGAGAGATGACGATTCAGCAGTTTGCAGCAGCTGTGG	36270
Oy	8195	TCACCAAGCGAAGCTATCTGCTGCGAGTAAATCTATGTTAGAGTTGAGGAAGGAATT	8254
Db	36271	CAGAGTAGTACAGCATCAGCAGTTTGCAGCAGCTGTGGCAGGAGAGTAGCAGCATCAGCAG	36330
Oy	8255	TGTTTCAGAACTAATTCGATTAATCAATTTCTAAATGTAAGGTTTGGATGAAATTAAG	8314
Db	36331	TTGCAGTAGCTGTGGCAGGAGAGATGACAGATTCAGAGTTTGCAGTAGCTGTGGCAGGAG	36390
Oy	8315	TAACTGCTAAATCTTCTGTAGTATCAGAAATGAGAGGAATTCGCGAGCAGAGATA	8374
Db	36391	GAATGACGATCAGTAGTTTCAATGCTGTGGCAGAGAGATGACAGCAGCAGCTGTGG	36450
Oy	8375	ATACTTTCACAGCAAAAGTAATCTGAATCCGTAATCCGTTTACGAAACGAATTAATG	8434
Db	36451	CAGTAGGTGTGGCAGGAGAGATGACAGCAGCAGTTTGCAGTAGCTGTGGCAGGAGAG	36510
Oy	8435	AAATATATGATTTACACAAAATAATATATTTCAGAAATCAATGCTCTTGTCTTTAAATGATA	8494
Db	36511	TAGCAGCAGCAGCAGTTGCGATGACTGTGGCAGGAGATGACAGCATTCGCGAGTTGCAGCAG	36570
Oy	8495	CAAAAGATGAAGCGAATATAGATCTTTAGCGGTAGCCGGTGTGATGCAACAAGAACAA	8554
Db	36571	CTGTGGCAGAGAGATGACAGCATTCGCGAGTTTGCAGCAGCTGTGGCAGGAGAGTAGCAG	36630
Oy	8555	ACAAAGCATTTACGAGA-----TCAACAAAGTTAACTTTCTAACCTGTAATGGAAGAA	8608

Db	Accession	Length	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	36631	CATCGCAGTTGGCAGCAGCTGTGGCAGGAGAGTACACATTCGCGAGTTGCAGCAGCTG	36930						
Qy	8609	ACGTATTCCTCACTTCCTGCGCAAAAAGCTTTGGCTATAAAAATGAAAATTATGAAATGTAAGAAG	8668						
Db	36691	TGGCAGCAGGAGATGACGACATTCGGCAGTTGCAGTACCTGTGGCAGGATGACGACGCGAG	36756						
Qy	8669	GAACCTGGAGGAGCCTTAGTGTGGAGCCGAAACAGCAGCCGTTGAAAAATTTACAAAGAGTA	8728						
Db	36751	CAGTTGCAGTAGCTGTGGCAGGAGGTGACAGCAGCAGCACTTGCAGTAGAGCTGTGGTAGAG	36810						
Qy	8729	CTACAGCAGCAGCTTGTGTTGCAGCAAAATTTGGGAAATTGGAGATAAATTGAAACGATTGC	8788						
Db	36811	TAGCAGCAGTACAGCAGTTGCAGTAGTGTGGCAGGAGAGTAGACAGCATCAGCAGTTGCAG	36870						
Qy	8789	GAGATTAATACGATTGTGAAGTCAACGAGACCGAAACAAAGAGGTCTTGTTCGGAAGA	8848						
Db	36871	TAGCTGTGGCAGGAGGAGTAGCAGCATCAGCAGTTGCAGTGTGGCAGGAGGAGTAG	36930						
Qy	8849	ATGCTATTTCTGTGAATAAATACAAATTTCCGGGGGAAACAAATCATCATTAAGATTAAG	8908						
Db	36931	CAGCAGCAGCAGGTTGCGTAGCTGTGGCAGGAGTACAGCATCAGTAGTTGCGTAGCTG	36999						
Qy	8909	CCAGATTTGTTGGAAACCGGAAAGTGAATGTAGATCTTTGATGAATCTGATGTAGATC	8968						
Db	36991	TGCGACAGCAGATGACGACATCAGCAGCTTGCAGTACTGTGTGGCAGGAGAGTAGCAGCAT	37050						
Qy	8969	TACAAAGAAAAAGTGTGTGCTATGTGGAAATTTGTAATTGGAATTTGATGTGA	9021						
Db	37051	CAGCAGTTGCGATGTGTGGCAGGAGAGTAGCAGCATCAGCAGTAGTTGCGAGTA	37103						
RESULT 29									
AC122227		208332 bp	DNA	linear	ROD 01-JAN-2004				
LOCUS									
DEFINITION		Mus musculus chromosome 12 clone RP23-131H10, complete sequence.							
ACCESSION		AC122227							
VERSION		AC122227.4							
KEYWORDS		GI:40539152							
SOURCE		HTG.							
ORGANISM		Mus musculus (house mouse)							
		Mus musculus							
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
		Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;							
		Sciurognathi; Muridae; Murinae; Mus.							
REFERENCE		1 (bases 1 to 208332)							
AUTHORS		Wilson, R.K.							
TITLE		The sequence of Mus musculus clone							
JOURNAL		Unpublished							
REFERENCE		2 (bases 1 to 208332)							
AUTHORS		McPherson, J.D. and Waterston, R.H.							
TITLE		Direct Submission							
JOURNAL		Submitted (22-MAY-2002) Genome Sequencing Center, 4444 Forest Park							
		Parkway, St. Louis, MO 63108, USA							
REFERENCE		3 (bases 1 to 208332)							
AUTHORS		McPherson, J.D. and Waterston, R.H.							
TITLE		Direct Submission							
JOURNAL		Submitted (28-FEB-2003) Genome Sequencing Center, 4444 Forest Park							
		Parkway, St. Louis, MO 63108, USA							
REFERENCE		4 (bases 1 to 208332)							
AUTHORS		Wilson, R.K.							
TITLE		Direct Submission							
JOURNAL		Submitted (01-JAN-2004) Genome Sequencing Center, 4444 Forest Park							
		Parkway, St. Louis, MO 63108, USA							
COMMENT		On Jan 1, 2004 this sequence version replaced gi:28604203.							


```
repeat_region /tpc_family="ERV1"
30980..31076
repeat_region /tpc_family="B2"
31015..31332
repeat_region /tpc_family="B4"
31579..31762
repeat_region /tpc_family="A1u"
31737..31803
repeat_region /tpc_family="B4"
31850..31977
repeat_region /tpc_family="B4"
32927..33021
repeat_region /tpc_family="B4"
35441..35733
repeat_region /tpc_family="L1"
35770..35980
repeat_region /tpc_family="B2"
36001..36281
repeat_region /tpc_family="L1"
36251..36378
repeat_region /tpc_family="L1"
36379..37961
unsure /note="Unresolved tandem repeat."
repeat_region /tpc_family="L1"
37960..38475
repeat_region /tpc_family="L1"
38670..39113
repeat_region /tpc_family="MaLR"
39210..39264
repeat_region /tpc_family="ERV1"
39623..39775
repeat_region /tpc_family="L1"
41596..41744
repeat_region /tpc_family="B4"
43065..43277
repeat_region /tpc_family="RMR6A"
43411..44127
repeat_region /tpc_family="L1"
44131..44614
repeat_region /tpc_family="L1"
```

Query Match 1.1%; Score 106.2; DB 9; Length 216959;

Best Local Similarity 42.3%; Pred. No. 2.5e-05; Matches 652; Conservative 0; Mismatches 888; Indels 3; Gaps 1;

```
6315 ATCAGGAGATTACCAATTCATCGAGTAGTAGGTTCTGTGGTGTGGAATGT 6374
37943 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37884
6375 AGGAGTAGAGCTTCTTCTGATACCAATTAATAAAGAAATACCAAGACAAAGTTGG 6434
37883 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37824
6435 AAAAACTACAATCTCGATGAAGTTTCGAGAAAGCTGAATTAACAGCAGATTCTTA 6494
37823 AGCTGAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37764
6495 GCAAGAATTTCTCTTTTGAAGTCGAGTCGAGCAGCCGGGGTAGAGCCGGAGTGC 6554
37763 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37704
6555 AGGAACCGTTTCGTAATCAATTTGCAAGAAAGACGAAAGTAGATGTGAAAGAACAA 6614
37703 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37644
6615 GATTTGGTAAAAAAGCTGAATTCACGAAAAAGTTATAGTTCTGTTGCAATTGGAAA 6674
37643 AGCTGAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37584
6675 TGCCGAGTCGAGTGGCTGCAAAAGAGCTGGAATTTGAGAGCAGTGGCAGTTACCA 6734
37583 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37524
6735 AGATGAATCAACACGAGACAGAGTAGAATAATCTAAATTAATGACTCGAAACAAAGTT 6794
```

```
37523 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37464
6795 AGATGTAATAGCAAAAAAGATTAATCAAGTACTGGATTCGGTTAGCCGAGCTGG 6854
37463 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37404
6855 AATCTTGCAGCCGAGATCTGAGTGGTTCTGTCAATAATTAATGCAAAATTAAGTAGA 6914
37403 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37344
6915 AACAGATTCGATCATAGTACTTTACCTCTTCTAATGATGTAAATGTAAGTCTTTAA 6974
37343 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37284
6975 TAAATTTTCGAATTCCTTGAACAGCCGGTGAAGAGCCGAGGTCTTGAAGCATTACCG 7034
37283 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37224
7035 AGTGTTTCTGTAACTAATAATAGTTCTGTGATAGCTCGAGTTCAATACTCTGA 7094
37223 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37164
7095 TTGACTTCGTTACGAGAAAAAGTAATGTAAACGCAAAAGAGAAAAATTAAGCA 7154
37163 AGCTGAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37104
7155 AACGACAGCAATGAGGAATTCGAGAGAGAGCAATCGAGCCAAATGTTGGTAAATA 7214
37103 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 37044
7215 TTTTGAACAGCTGTAGAAATAGAAAAATTTCTGAAGAAAAAGAACAGAAATTTTAA 7274
37043 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36984
7275 AACTTAGACGAA--GTTAACAAAGAACAGATTAATAAGTAATGATCTACGAAAA 7331
36983 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36924
7332 AATCTTAATTCAGAGGATTTCTACAGAAAGATCTCTGTAAGAGCGAATAGAGAGA 7391
36923 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36864
7392 TACTCAGGAGAAAGGAATTTAAAGCCATTGTGAAGCTTCTGATATATTGAAAAATGT 7451
36863 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36804
7452 AGATATTCAACAGAGACAAAGATTAATCACTTCTAATGAGTTGGAACTGCAG 7511
36803 AGCTGAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36744
7512 TCTTGCTCCGATCAGGAACAGTGGCAGTTACAAATTTTAAAGAAATTCGAGTTAC 7571
36743 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36684
7572 TGTGAAAAATCTTTTGAAGACAGCTGAAGAAAGTAATGTTAGATCGAATATACAG 7631
36683 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36624
7632 AAATGTTCTTTAACAGCATATCAAGGTCCTGTAGAGCAATTGGAAATAGAGCTGCCTA 7691
36623 AGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36564
7692 TGCAGAAATTAATTTAATGAAAGTCAAAATATCATGATTTAAATTTCTAAGCTTTTGG 7751
36563 AGCTGTACAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36504
7752 AAAAAATATTAATGTTATTGTAAGAGATAATTCGAAATTTGAGAGCGAAGCAAAAGATT 7811
36503 AGCTGTACAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGTAGCAGT 36444
7812 AACCGTAGAGCGTAGCTGCGGAGCAATTAATCTCAAAAGCA 7854
```


EDKALNEKIMSEEEVIEDONHEVPAEVSUTEI PGVEETTNESHVKGENVEN
VVDNSVDEIQFDDSSIEITYDSKOVFHEENYDSREPREDENIH SRKNEN
PVKIKENESVGVHPTTEYSKENIAEVPESHIMPEVSEPEVSEVIEEVEF
BEVTSKENIEIEAPGDIENSENIEEVPVTSBEI VONDNINTNTIDHFDLDD
DLIKIPEAORUKLRTLVDLVDLVEIKORDEKKNRDFLYICLRINKSEPTLM
KEYGSDSNKSSANDNTVYVIKIKOKKGRIVDTIANSLSYSLVNLITBEMKD
LINTTETSNKSKOHMKUDMKWKEBKFELFKRDLKIDMKRTYSIDONBENESM
NELSDELKNNTSDMNRNEQKLEKNEKESNLSKOLGIEKXIIMKNIYKVNINIE
NMDEWPKNVNNEKEKNDVSNLOENSBASVENYISDVDSKDMTNSINDSEKX
IVASSKTNEEYIILSRKDLINIIIVHMMVLDPKFDELKXKRFUNSIDKFI
KEKKIKDETVLDYIIDIIKRIIVEDTSHVNNIKKAIIDHTSHDMFRLROKNAON
SISQVDILKEKXYKILISFEDNKNNNNNDEVPKELKNIKCEBNVSPILRSNK
KQDPEEKKNKNDLNTLISDSGKININIBESKNGNDKNSGNNSRKRDR
RYERKEKODLRFVLOLDYENIFEOIONMENONKOKNKETNIIKTSIDIKNLLRE
YKKEKNIINOSKEFSVNNEN"

ORIGIN

Query Match 1.1%; Score 105.6; DB 2; Length 16377;
Best Local Similarity 41.8%; Pred. No. 7.9e-05;
Matches 1838; Conservative 0; Mismatches 2479; Indels 75; Gaps 17;

QY 5131 GATGAACGGAAGCTTAGTGTGATCTGAGTTGAAGAGTAAGTCTTTCATGTA 5190
DB 3325 GTTGTGAAAAAGATCAGATCTGAAATTGTTGTAAGAAAGATCAGCTA 3384
QY 5191 GATGCAAAAGATCAAAAAACAATAATACATTCGCCGAATG---CAATGAGAGAAA 5247
DB 3385 GTTGTGAAAGAAAGATCAGCTA 3444
QY 5248 GGGGCTGGAGTTGAGACAAAGTTGCTCAAAATATTGGAACAATCACTTATAGT 5307
DB 3445 ATGTGAAAGAGAAAGATCAGCTA 3504
QY 5308 ATTGTAAAAACAGTAAATTAACAAGCGGAATGATCAGATGAAAAATATCAATG 5367
DB 3505 AATGTTCAAAAGTCAAAAGAGATTAATTGAAAAAGTAAGAAATATATTCTGTA 3564
QY 5368 ACTGCAAAAGATTATCTATGACCAATCTATAGCAGTCGAGATTGAGAGACAAAAG 5427
DB 3565 GCTAAGAGATGAATCTATGATGAAAAAATATACATTAGAAATCATCTGTTGATG 3624
QY 5428 GCCTCTGTCAGAGAGCTTCTGCACTACCTTGAATTAAGACAGTTTCTCATGTT 5487
DB 3625 GAGTCTTTTAAAGACAAGAGCTACAGATTAATGATCAGTCAAGACGTAATTCG 3684
QY 5488 GATCAAACTGATATTGACAAAGATTAGAGAAAGAAATATGAAATTAAGAGAAAG 5547
DB 3685 GATATTATTATTAACATCTCCATTAGAGAAAGCAAAATTTGAATTACCAACAG 3744
QY 5548 AATGTTAATGTTCTAGCTGAAAAATACGA---GTCAAGTGTCACAATGCGACAGTCT 5604
DB 3745 TATATTACTAGCTGATATACACCTAGAGTGAAGAAACAGAAATATATCTAATA 3804
QY 5605 TCCGAGACAGTGCAGACAGCTGAGAGCTGAGTACGATTATATTAATACACA 5664
DB 3805 TACGAAAAATATATGAAAGAGTTGAAAAAGCAAAATTTCAAGAGAAAAATTA 3864
QY 5665 AATACTTTCGACATATAAAAAATGACTCAAAATGTACGAAATGCTTGTGTA 5724
DB 3865 AATATTCCGTAAGAAAGAAATCAGTTACAGAAAGATATAGTACAAACCATCATTA 3924
QY 5725 AATCTCATTCATCTATTAACAATTTGAAATTTGAGCTGAGATTGAGACTGAGAG 5784
DB 3925 CAGGAAGTTGAACAAAAAGATCAGATCTAATGAAATTTGAGAAATCAATTA 3984
QY 5785 GGAAGTACAGGTTCTGTAGAGTGAATAAGATTGAATTAATTCAGTACAGAA 5844
DB 3985 GAAAAAATATATACGAAAGATCTGTTA 4044
QY 5845 CATGCAAAATCACTGCGAAGGAAATGTCGAGTTATTACAGAGTCTGATCGGTA 5904
DB 4045 ACTGAAGAGCTCTTAGAAGAGATGAATCAATTA 4104

QY 5905 GCTAATTATGACAGAACAGTCTGAGAGTGCCCGGTGACAGCAATAGAGCCTCAAC 5964
DB 4105 TCAGCTTGTGAAGAAATAGTAGAGAGAGAAATCATCTAGTGAAGAAATAGTAGAG 4164
QY 5965 GTGAATGAATTTACAGATCTACAAAGCATATGTAAAGATTCTA---CAGTGTGCT 6021
DB 4165 GAAAGATCATCTAGTGAAGAAATAGTAGAGAGAGAAATCATCTAGTGAAGAAATAG 4224
QY 6022 AAAAGAAAGAACAGATGATTAATTAATTAATCTACAGAGGCAAGTAGATTAAG 6081
DB 4225 GAGAGAGAGAAATCATCTAGTGAAGAAATAGTAGAGAGAGAAATCATCTAGTAGAG 4284
QY 6082 GTATTCAAAATCTTAATATTAACAGAGACTTATCACAAAAAGAAAAATAGTAT 6141
DB 4285 ATATATAGAAAGAGTCAAGTACAGTAAGAAAGTCTTAGAAGAAAGAGATCGGTA 4344
QY 6142 AAAAGATTTGTTACCAATAGTTCAGCTACTCATCTTAATCTTTATTGGCAAAATCC 6201
DB 4345 GAAATGTGAGAGAAAGAGATCACTA 4404
QY 6202 GCTGGTTCAAGACAAAGCCGAGTGACAGAACTGTTAATATCAACAAGTTTATG 6258
DB 4405 GAAAGAGTATTAGAAAGAAAGAGATTGCTACTGAAAGATTGTAAGACCAACAGATCT 4464
QY 6259 GAAACAGAGCTCTTGTGAAAAATTTCTATATTAATGCAAAACATTAATCTGTA 6318
DB 4465 GTTATTGAAGAAATGTGAAACAGAAATCATCTAGTGAAGAAATCTAGAAAGT 4524
QY 6319 GAGATTAACAGAAATTCATGAGAGTATAGTCTGTTGTTGTTG---GTGGAATGTA 6375
DB 4525 TCAATGTTGAAGAAATGTACAGAGAAAGGTCCTGTTGCTCAAGAAATGTACATGA 4584
QY 6376 GAGTAGAGAGCTTCTTCTGATACCAATTTATTAAGAAATATACCAAGACAGAGTTG 6435
DB 4585 GAAATATCACTACGAAAAACATGATGAGTGAATCACTACGAAAGAAATTTGTA 4644
QY 6436 AAAC-----TACAATGCTGATGAGAGTTTCGAGAGAGCTGAAATTAACAGAG 6489
DB 4645 GAAAAAGTTGATCAGTTCTTGAAGAAATTAATAGAGAAAGAGTGTGAGTCTGA 4704
QY 6480 TCTAAGCAAGAAATTTCTCTTTTGGAGTCCGAGTCCGACAGCCGGGTAGAGCCGA 6549
DB 4705 ATAGTAGAAGAGATCAGTTACTGAAGAAAGTGTAGAGAGAGAAATTAATTA 4764
QY 6550 GTGCAAGAACGTTTCGTAATCAATTTTCAGAGAAAGACGAGATGATGTGAAAG 6609
DB 4765 GTTGTGAAGCCGAATCGGTTACTGAGAAATAGTACAAAAGAGATTCAGATGCTGA 4824
QY 6610 GCAAGATTTTGTAAAA-----AAAGCTGAGATTACGCAAAACGTTATAGTTCT 6660
DB 4825 GAAATTTTAGACAAAGAGATCAATGATGAAGAAATATTAGAAAGAAATCAATGTT 4884
QY 6661 GTTGCATTTGGAATATGCGCAGTCCGAGTGGCTGCAAAAGAGCTGGAATTTGAGCA 6717
DB 4885 GAAAGAAATTTGTGACCGAAGAGATCTGTTACTGAAAGAAATAGTATTAACGAT 4944
QY 6718 GCAAGTGCAGTTTACCAAGATGAATCAACAGAGCAAGAGTGAAGAAATTTCTAAAT 6777
DB 4945 GCTGAAGAGTCAAGAAAGAAAGAAATTAAGTACAGAGAAAGCTGTACAGTATGAAG 5004
QY 6778 ATGACTCGAAAAAAGTATGATGTAATAGCAGAAAAAGATTAATTCAGATCTGGAATC 6837
DB 5005 GTTACTGAAGAAATTAAGAAAGAAAGAAATCGATTACTGAGAAATGAAGAAATCA 5064
QY 6838 GGTTCAGCCGAGCTGGAATTTCTGACCCGAGATCTGAGAGTGTGTTCTGTCAATAT 6897
DB 5065 GCTTTTGAAGAAATTAATGAAAGAAAGAAAGACCAATTAATGCGAAATTTGAAGAA 5124
QY 6898 ATTGCAATTAAGTGAAGACAGATTCAGATCAATGATCTTTTACTGTAAGTGA 6957
DB 5125 GATTAAGTACTGAGAAATATGATGAATGAATATAGTTACTGAGAGAGTCA 5184
QY 6958 AATGTAAAAAGCTCTTAATTAATTTGAAATTTCTTGAAGCCGGTGAAGAGCCGAG 7017

Db 5185 GAAAAATTAGTTAATGAACAAATTGTAGCTGAGAGAAATCTGTACTGAAAAACCACTA 5244
Qy 7018 CTTGCAGCAGTTTACCGAGTGGTTCTGTGTTAACTATTAATTAAGTTCTGTATAGCTCGA 7077
Db 5245 GAAGTTGAACGATCAGCTACTGAAAGATCTTGTGAAAGAGGACATCAGTACTGAAAAA 5304
Qy 7078 GTTCACAATTACTGTATTTGACTTCGCTAGCAGAAAAAGTAAATGTACCGCAAAAGAG 7137
Db 5305 GTTTCAGTGCATGAAGATCAACTACTGAAAC---AAATACTAGAAATCAAGTTGCTTAA 5361
Qy 7118 GAAAAAATATTAAACCAACGACCAAAATGACAGGAATCGGAGGACGACATCGGAGCC 7197
Db 5362 GAAATAGTAGAGAAAGAAAGTATCTGTGATGACAAATTAATAGAGAAAGAAATCTGTT 5421
Qy 7198 AATGCTTGTAATTAATTTTGGAACTGTGAGAGATAGAAAAATTTCTGAAGAAAA 7257
Db 5422 GATGAAGTTGTGAAGAAAGAAAGATCAGTTATTTGAAGAGATGTGAGAAAGAAAGATCA 5481
Qy 7258 GGAACAGAAATTTTAAAACTTTAGACGAAGTTAACAAAGAACAGATAAAAAGTAAAT 7317
Db 5482 GTTCTGAAGAAATTTGGAAAGAAACTATCAGGTTCTGAAGAAATTTGAAAGACGA 5541
Qy 7318 GATGCTAGAAAAAATCTTAAATCAGAGGATATTTCTACAGAAATATCTCTGTAATA 7377
Db 5542 TGGGTTACTGATGCTTTTATGAGACAAAGAGATCTGTTATGAAGAAATTTGAAGAAAT 5601
Qy 7378 GCGGATAGAGGATATCTCAGGAGAGAAAGAAATTAAGCCATGTGAGAACTTGTATAT 7437
Db 5602 GTAGATGGAGAGATCATTACGGAAGAAATTTGAAAGATGATCAGCTAATGAAAA 5661
Qy 7438 ATTGAAAAAATGTAGATATTAACAAGAGACAAAGATTAATATCATCTTCTAGTGTGT 7497
Db 5662 AATAGTAGAGAAAGAACCATCTGTGTGAAGATTTTAGAAAAAGAAAGATTTGTAT 5721
Qy 7498 TTGGGAACTGCAGGCTTCTCCGATCAGAAACAGTGCAGCTTCAAAATTTAAAAA 7557
Db 5722 GAAGAAATTAATGAAGAAAGGTCGTAATGAACAAGTTGAAGATCA-----AAAAA 5775
Qy 7558 AATTCGGAGTTACTGTGAAAAATCTTTGTGAAGCAGCTGAAGAAATTAATGTTAGA 7617
Db 5776 GTTTCGAAAAATCAGAAAGATCATAGCTATTGAAAGATTTAAAGAAATTAAGAAAG 5835
Qy 7618 TCGGATATTACAGAAATGTGCTTTACAGCATATCAAGTCTCTAGAGAGATTTGGA 7677
Db 5836 GAATCATTAATGAAABAAATAGTAAAGAAAGAAAGATCTGTTACTGAAGAAATGTACGA 5895
Qy 7678 ATAGAGCTGCTTATGCAAGATTAATCTTAATGAAGATCAAAATATCAGTATTTAAAT 7737
Db 5896 CAAGAGAAATCTACTACTGAAAAAATAGTAAAGATGATCA-----CCTACTGAAGAT 5949
Qy 7738 TCTAAGCTATTAGGAAAAAATATGATGTTATTTGTAAGATTAATCGGAATGAGAGG 7997
Db 5950 TTTGTGAACAACAATTCAGTTACTGAAAAAGTTTAGAACAGAGAGATCAACACT 6009
Qy 7798 GAAGCAAAAGGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTAATCTCAAAAGCAAG 7857
Db 6010 GAAGTTGCAGAAACGTTGAAGAAAAAGATTCGCTTGTGATGAACATGAACAAGAAAT 6069
Qy 7858 AATGAATGAATTCAGAGGTGAATTTGAGAAAGATTTTCAATGAAGAA--AATAGAG 7915
Db 6070 GTATCAGTAAATGACAAAGTTTACATGTGAGAAAAAATCACTTACTAAAGAAATAGTGAC 6129
Qy 7916 TAACTAGCCTTTAAAGAAATTTGGAAGAAA-----TCAATGCAAAAGTGGA 7965
Db 6130 GAAGTGTCTGTACTGAAGAAATAGTAGAGAAATGATCAGTTACGGAAGAGATGAC 6189
Qy 7966 AAAGAAAACAGAGTACTGCTGAATCTCAAGAGCTTCTAGAGAGAGTACGAGGCGCA 8025
Db 6190 GAAACAGAGATCCGTTACTGAAGAAATTAAGAAAGCAACAGTAACTGAAGAAAGTGTGA 6249
Qy 8026 GGAATTAATTTCCGAAGCAAAAGATGCCGGAAGCTTTATTGAAAGTTAGTACAAAAATCC 8085

Db 6250 GAAGATGATCAGTAATCTGAAGAAAGTGTAGAAAGATGATCATGTTATTCAAGAACTGTA 6309
Qy 8086 GGAAGAAATTTTTCATGACAGATATATGAAATATGGAAGCAACATTAATTAAGTA 8145
Db 6310 GAAGATGATCAGTTACTGAAGAAATAGTACAGAAAAATGATCAGTTACTGAAGAAAT 6369
Qy 8146 ACAGCAGTTTCTAAAGCAGTAACAGTTCGTATTTGGAGAGATTGAGTCCAAAGCA 8205
Db 6370 GTAGAGAAAGAGATCAGTTAATGAAGAAAGTTGAAGAAAGATCAGTTCTTGGAGAA 6429
Qy 8206 GAACTACTGTCGACAGTAAACTATGTGTGAAGTTGAGAGAAAGAAATTTGTCAAGCA 8265
Db 6430 GTTATGAAGACAGATATGTTACTGAAGAGTGAAGAAAGAGATCAGTTGTGAAGAA 6489
Qy 8266 AATGATTTGAATGCAATTTCTAAAGTGAAGGTTTGAATGAAGATTAAGTAAGTCT--- 8322
Db 6490 ATTGTAGAGAGAAAGATCAGTTGTGAAGAAATTTGAGAGAGAAAGATCAGTTGTT 6549
Qy 8323 -----AATCTTGTAGTATCAGAAATCGAGAGAAATTCGCGAGCAGAGTGAATA 8377
Db 6550 GAAGAAATTTGAGAGAGAGAAAGATCAGTTGTGAAGAAATTTGAGAGAGAGAGATCA 6609
Qy 8378 CTTCTACAGCAAAAGTAATATCTGAATCCGATGTTTACGAAAGCAAGATTTAGAA 8437
Db 6610 GTTGTGAAGAAATTTGTAGAGAGAGAGATCAGTTGTGAAGAAATAGTAGAGAGGA 6669
Qy 8438 AATATGATTACAAAAAATATATTTCA--GAGTCAATGCTCTTGTAAATGATACA 8496
Db 6670 GATCAGTTGTGAAGAAATTTGAGAGAGAGAGATCAGTTGTGAAGAAATTTAGAG 6729
Qy 8497 AAGATGAAGGAATATGAATCTTTAGCGGTAGCCGCTGTGATGCACAAAGAACAAAC 8556
Db 6730 GAGAGAGATCAGTTGTGAAGAAATTTAGAGAGAGAGATCAGTTGTGAAGAAAT 6789
Qy 8557 AAAGATTTAGCAGATCAACAAAGTTAACTTCTACACTGAATAGAGAGAAACGTATCT 8616
Db 6790 GTAGAGAGAGAGATCAGTTGTGAAGAAATTTGAGAGAGAGAGATCAGTTGTGA 6849
Qy 8617 CAATCTCGTCAAAAGCTTTGCTTAAAAATGAAATTAATGAATGTAAAGAACTGGA 8676
Db 6850 GAAATTTGAGAGAGAGAGATCAGTTGTGAAGAAATTTGAGAGAGAGAGATCAGTT 6909
Qy 8677 GAGCCCTTATGTCGAGCGGAAACAGAGCCGTGAAATTAACAAGAGTACTACAGA 8736
Db 6910 GTTGAAGAAATTTGAGAGAGAGAGATCAGTTGTGAAGAAATTTGAGAGAGAGAG 6969
Qy 8737 GCATGTTTGCAGAAATTTGGAATTTGAGATTAATTAAGAAACGATTCAGAGATAAT 8796
Db 6970 TCAAGTTTGAAGAAATTT---GTAGAGAGAGAGATCAGTTTCTGAAGTATGTTGA 7026
Qy 8797 ACAGTTTGAAGATCAACGAGACGAAACAAAGAGGCTTTGTGGAAGAAATGATTT 8856
Db 7027 ACTGATTTAGTAAATGATGAATATAGTAAACAAGCTCCATTTACTGAAGAGTTGAAGA 7086
Qy 8857 TCTGTGAAGAAATACATTTCAAGGGAACAATATCATCATTTGAGAGATAAGCAGAAAT 8916
Db 7087 CAAGTATCAGTTAAGCATAAATTAATAGAAAGTCACTGTTGCTAACAAGTTGAAGA 7146
Qy 8917 GTTGAACCGGAAGTGAATATGATGCTTTGAATGAATGAACTGATGATGATCAAGGA 8976
Db 7147 TCGAATCTAATTAATCAAGTTTCCCAAGAGAAAGAAACGAAAGAAAGATTTGTTAT 7206
Qy 8977 AAAAGTGTGCTATAGTGAATTTGATTTGAATGAAATGTTGATTAATATGATTAAG 9036
Db 7207 GAAAAAGTTGAAGAAACAGAGCGGTTACTGAAGAAATAGTTCAAGATGATTAATTAACA 7266
Qy 9037 AAAATGTGAAGCCAAATTCGAGAGACATGCTATTTGAAACTCTGGAAGAAAGAA 9096
Db 7267 GAAGAAATTTTGAAGAAATCTGAATCCGTTAATGGGAAATTAATTAACAAGATCAGAT 7326
Qy 9097 TATCAAGATTTACAGAGCAAAAGTAATATTTCTTGAAGAAAGAGACCTGAGCTGCA 9156
Db 7327 GCTGAGAAATTTTGAAGACAGATTTCTTACTGAAGAAAGTTGTAGGTCAAGCAGATCA 7286

```
QY 9157 GCTGCAATATCGAATGTACACATTTCCAAATGAGATGATATTAATAATTTGGCAAGCAG 9216
Db 7387 ACTTCTGAAAGAAATAGTAAAGAGAGATCAGTTACTTAAAGAGTTAGAGAAAAGGA 7446
QY 9217 TATGATCTTCTCAATTAATTAACCAAAATTCAAAAATATATTACTTTAGCATCAAGT 9276
Db 7447 TCAGTAACTGAAAGAAATTTAGTAGAGATCAGTAACCTGAAGAAATTTAGATGAAGGA 7506
QY 9277 AATGAAATCGAATGTGAATGT---TCATGGGGTGGCTGAAGCAAGAGGTCCAGAGCCAA 9333
Db 7507 TCAGTTACTGAAAGAGTTGTAGAACAAAGGGGTTCAATTGCTCAAGAAATAGTAGAAGGA 7566
QY 9334 GCGACAGTTAGTGTAAAGATCAAAATTAATAGAACTAATATGTTGATTTAGCAGGAAA 9393
Db 7567 GAATCAGCTACTGAAAGAAATTAATCGGATGAAATCTAATGTGAAGAGTCTTGAAGAAA 7626
QY 9394 ATTTAAACAGAGGAGAAACATCAATGATATATGCCGATATGATTAATAATTAATAGT 9453
Db 7627 GAAGATCTGCTACTGAAAGAAATAGTACAAAGATGATCAGTACTAATGATTTTGTAGGA 7686
QY 9454 AAGACAAATTC 9465
Db 7687 AACAAGGATCT 7698

RESULT 32
AC129389_1/c
WCOMMENT
Sequence split into 5 fragments LOCUS AC129389 Accession AC129389
Fragment Name Begin End
AC129389_0 1 110000
AC129389_1 100001 210000
AC129389_2 200001 310000
AC129389_3 300001 410000
AC129389_4 400001 425691
Continuation (2 of 5) of AC129389 from base 100001 (AC129389 Ratius norvegicus clone CH2)

Query Match 1.1%; Score 105.2; DB 14; Length 110000;
Beet Local Similarity 42.1%; Pred. No. 4,4e-05;
Matches 591; Conservative 0; Mismatches 806; Indels 7; Gaps 3;

QY 7105 GTACAGAAAAAGTAAATGTAAACGCAAAAGAGAAAAATATTAAGCAACACACACA 7164
Db 16649 GAAGAGAAAGAAAGAAAGAAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16530
QY 7165 AATGACGAATTCGAGAGACGCAATCGAGCCAATGCTTGTGTAATTAATTTTGGACA 7224
Db 16589 GAAGACGAAGAAAGAAAGAAAGAAAGAGAAAGAAAGAAAGAAAGCAAGAAAGAAAGAA 16530
QY 7225 GCTGTAGAAGATAGAAAAATTTCTGAAGAAAAGAACAGAGTTTAAATCTTAAGC 7284
Db 16529 GAAGCAGCAAGAAAGCAAGACAGAACAGAAAGCAAGACAGAAAGCAGAAAGAA 16470
QY 7285 GAAGTTAACAAAGAACAGATTAATAAAGTAAATGATGCTACGAAAAAATCTTAACATCA 7344
Db 16469 GAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16410
QY 7345 GCAGGTATTTCTACAGAGATATCTTCTGTAAAGCGATAGAGAGATCTCAGGAGAA 7404
Db 16409 GAAGAAAGAAAGAAAGAAAGAAAGAAAGCAAGAAAGAAAGAAAGAAAGAAAGAA 16350
QY 7405 GGAATTAAGCATTTGTGAAGCTTCTGATATTAATTGAAAAATGTGATATTAACAACA 7464
Db 16349 GAAGAAAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAAGAAAGAAAGAG 16290
QY 7465 GAGGACAAGATTAATATCACTTCTACTGGTGTGGAACTGCAAGCTCTTGCCGCA 7524
Db 16289 AAGAAAGAAAGAAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16231
QY 7525 TCAGCAACAGTGGCAGTTACAAATATTAAGAAATTCGAGAGTTACTGTTGAAATCT 7584
Db 16230 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 16171
```

```
QY 7585 TTTGTAAAGCAGCTGAAGAAAAAGTAAATGTAGATCGATATTTACAGAAATGTCCTTTA 7644
Db 16170 NNNNNNNGAAGAAAGAAAGAAAGAAAGAAAGAAAGCAAGAAAGAAAGCTTTTGAAGA 16111
QY 7645 ACAGCATATCAAGGTCTGTAGAGACATTTGGAAATAGAGAGTGTCTATGACGAATTAAT 7704
Db 16110 AGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16051
QY 7705 TCTAATGGAAGATCAATATATCAGTATTAATAATTTCAAGCTATTAGAAAAATTTGAT 7764
Db 16050 GAAGAAAGAAAGCAGCAGAAAGAAAGAAAGAAAGAAAGAAAGCAGAAAGCAGAA 15991
QY 7765 GTTATTTGTTAAAGATTAATTCGGAATTGAGACCGAAGCAAAAGCTTAACCTAGAGCG 7824
Db 15990 GCAGAAAGCAGAAAGCAGAAAGCAGAAAGCAGAAAGAAAGAAAGAAAGAAAGAA 15931
QY 7825 GTAGCTGCCGAGCCATTATCTCAAAAGCAAGAAAGAAATGAAATTCAGAGGTTGAAAT 7884
Db 15930 GAAGAAAGAAAGAAAGAAAGATTTTGTAGCAGAAAGAAAGAAAGAAAGAAAGAA 15871
QY 7885 GAGAAAGTATTTTCAATGAAGAAATAGAGTAACTAGCCCTTTAAAGCAATTTGAAAG 7944
Db 15870 GCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAAGAAAGAAAGAAAGAA 15811
QY 7945 GAATCAATGTCAAAAGTGAAAAAGAAAAACAGAGTGACTGCTGAATCTCAAGAGCTTCT 8004
Db 15810 GAAGAAAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 15751
QY 8005 GTAGAGCAGATAGCAGAGGGCAGAAATTTTCCGAAGCAAAAGATGCCGAAAGCTTAT 8064
Db 15750 GAAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAAGAAAGAAAGAAAGAA 15692
QY 8065 TTGAAAGTTAGTACAAATTCGGAAGAAAGTATTTTTCATGCAATATGTGAATATGGA 8124
Db 15691 --AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAAG 15635
QY 8125 GCACACATTAATTAATGAAGTAAACAGCAGTTTCTTAAGCAGATTAACAGGTTCTGTAT 8184
Db 15634 GAAGAAAGAAAGCAGAAAGAAAGAAAGAAAGCAAGAAAGCAGAAAGCAGAA 15575
QY 8185 GAGTGTGAGTCACCAAGGCAGAAAGCTACTGCTCAGGTTAAATCTATGTTGAAGTTAG 8244
Db 15574 GCAGAAAGAAAGAAAGAAAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 15515
QY 8245 GAAGAAATTTGTTCAGAAACAAATCGATGGAATCAATTTCTAAAGTGAAGGTTTGAT 8304
Db 15514 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 15455
QY 8305 GAAGATTAAGTAACTGCTAAATCTTCTGTAGTATCAGAAATGAGAGGAATTTGCCGA 8364
Db 15454 GAAGAAAGAAAGCAGCAGAAAGAAAGCAGAAAGCAGAAAGAAAGCAGAAAGAA 15395
QY 8365 GCAGAGTGAATTAATCTTA--CAGCAAAAGTAAATTAATCTGAATCCGTGTTTGAAG 8422
Db 15394 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 15335
QY 8423 AGCAAGTTATGAATAATATGATTAACAAAAAATATATTTCAAGAGTCAATGCTCTTG 8482
Db 15334 AAGAAAGAAAGAAAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAGAA 15275
QY 8483 CTTTAAATGATCAAGAAATGAAG 8506
Db 15274 CAGAAAGAAAGAAAGAAAGAAAG 15251
```

```
RESULT 33
E10125
LOCUS E10125 2069 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA encoding an immunogenicity protein.
ACCESSION E10125
VERSION E10125.1 GI:22026753
KEYWORDS UP 1995284392-R/I.
```

SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2069)
AUTHORS Doi, H., Nagakuchi, Y., Tanaka, Y. and Fujiseki, Y.
TITLE GENE CLONE MANIFESTING CHICKEN LEUCOCYTOZON IMMUNOGENIC PROTEIN
AND GENE RECOMBINANT VACCINE AGAINST CHICKEN LEUCOCYTOZON
JOURNAL Patent: JP 1995284392-A 1 31-OCT-1995;
DOUBUTSUYOU SEIBUTSUGAKUTEKI SEIZAI KYOKAI, KITASATO INST:THE
COMMENT OS Leucocytozon caulleryi
PN JP 1995284392-A/1
PD 31-OCT-1995
PF 19-APR-1994 JP 1994080643
PI DOI HIROHITO, NAGAKUCHI YOSHIO, TANAKA YOSHIO, FUJISAKI YUUIRO
PC C12N15/09,A61K39/015,C12P21/02;
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..2069
FT /strain='shizuoka'
FT /organism='Leucocytozon caulleryi' FT
FT misc_feature 1..25
FT /note='EcORI adaptor'
FT misc_feature 2045..2069
FT /note='EcORI adaptor'
FT misc_feature 25..2046
FT /product='immunogenicity protein'.
FEATURES
source Location/Qualifiers
1..2069
/organism="unidentified"
/mol_type="genomic RNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 1.1%; Score 104.2; DB 6; Length 2069;
Best Local Similarity 43.3%; Pred. No. 0.00026;
Matches 748; Conservative 0; Mismatches 963; Indels 16; Gaps 5;
QY 7110 AGAAAAAGTAAATGTAAACGCAAAAAGGAAAAAATATTAAAGCAAAACAGCAACAATGC 7169
DB 312 AAACACACATCTGAACACGAAATTAACATACCTAATGAAGAAAGATTAACAAATGA 371
QY 7170 AGGAATCGAGAGCAGCAATCGAGCCATGCTTGTGAATTAATTTTGAACACGCTGT 7229
DB 372 TGAHAACGAAACAACATTAATGAAGAAATTAATTAATAGTAGATGATTTTACAG 431
QY 7230 AGAAGTAAAGAAAAATTTGGAAGAAAAAGAACAGAGTTTAAAACTTAAACAGACGT 7289
DB 432 AGTAAATGTGAAGAAATCAGATGAAACAGAACATTAATGAATGAATGAATGAATGA 491
QY 7290 TAAACAAGAACAGATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 7349
DB 492 ACAAGAGACATGAHAAGTAACATGAAGAAAGAAAGAAAGTAACATGAAGA 551
QY 7350 TATTTCTACAGAGATCTTCTGTAAAAAGCGAATAGAGAGTACTCAGGAGAGAGAT 7409
DB 552 AATGAAAAAAGAGAGATGAAGATGAATGAATGAATGAAGAAAGAAAGTAACACA 611
QY 7410 TAAAGCATTTGTGAAGCTTCTGTATTAATTTGAAAAAATGTGATTTTAAACAGAGA 7469
DB 612 TGAAGAAATGAAGAAAGAGAGATGAAGAAATTAATGAAGAAAGAAAGAAAGAGT 671
QY 7470 CAAGAATTAATATCACTCTACTGGTGTGGGAATGACAGTCTTGCTTCCGCATCAG 7529
DB 672 AACACATGAAGAAAAAGAAAGAGCATGAAGAGTAATACATGAAGAAAGAAAGAA 751
QY 7530 AACAGTGGCAGTT--ACAAATATTAAAAAGAAATTCGGAAGTAACTGTGAAAAATCTTTT 7587
DB 732 AGAAGTAAACATGAAGAAATTAAGAAAAAGAGCAAGAAAGTAATACATGAAGAGA 791
QY 7588 GTGAAGACACTGAAAAAGTAAATGTTAGTTCGATTTACAGGAATGTTGCTTTAA 7647

DB 792 AAAAGAAAGTAAACATGAAGAAATAGAAAAAGAGCATGAAGAAATACATGA 851
QY 7648 GCATTCAGAGTCCCTGTAGAGCATTTGGGAATAGAGCTGCTATGACGAATTAATCT 7707
DB 852 GGAAGAAAAAGAGAAATGACATGAAGAAATGAAGAAAGAGCATGAAGAGTAAT 911
QY 7708 AATGAGATCAAAATATCATGATTTTAAATTTCTAAGTATTTAGAAAAAATATTGAT 7767
DB 912 ACATGAAGAAAGAAAGAAAGAGTAACATGAAG--AAATGAAAAAGAGCATGA 969
QY 7768 ATTGTAAAGATTAATCGAATTGAGACCGAAGAAAGATTAACCGTAGACCGGTA 7827
DB 970 GAAGTATATCATGAAGAAAGAAAGAAAGATTAACATGAAGAAAGAAAGAAAG 1029
QY 7828 GCTGCCGAGCCATTAATCTCAAAACCAAGATGAATGAATTCAGAGTTGAATTGAG 7887
DB 1030 CATGAAGAGTAAATCATGAAGAAAGAAAGAGTAAACATGAAGAAAGAAAGAA 1089
QY 7888 AAGAGTATTTCAATGAAGAAATAGAGTAACTAGCCCTTTAAAGAAATGGAAGAGA 7947
DB 1090 AAAGTAAACATGAAGAAAGAGAAAGAAAGATTAATCATGAAGAAAGAAAGAAAGTA 1149
QY 7948 ATCAATGTCAAGTGAAGAAAGAAAGAAAGAGTGAATCTGAAGAGCTTCTGTA 8007
DB 1150 ATACATGAAGAAAGAAAGAAAGAGTAAACATGAAGAAAGAAAGAAAGTAAACAT 1209
QY 8008 GGAGAGTGAAGAGGCGAGGAATTTTCCGAAGCAAAAGATCCGGAAGCTTTA---- 8063
DB 1210 GAAGAAAGAAAGAAAGAGTAAACATGAAGAAAGAAAGAAAGTAAACATGAAGAA 1269
QY 8064 -TTTGAAGTTAGTCAAAATCCGAGAGATTTTTCATGCA-GATTAATGAAATAG 8121
DB 1270 GAAAAAGAAAGTAAACATGAAGAAAGAAAGTAAACATGAAGAAAGAAAGAA 1329
QY 8122 GAAGCAACATTAATGAAGTAAACAGAGTTTCTAAAGCAGTAAACAGTCTCTGTA--- 8178
DB 1330 GAAGTAAACATGAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAGAAAGTAA 1389
QY 8179 ---TTGGAGAGTTGAGTCAACAGAGCAAGAGTAACTGCTGAGGTAAATGTA 8235
DB 1390 CATGAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAGTAAATGAAGAAAG 1449
QY 8236 GAAGTTGAAGAGAAATTTGTTCAAGCAATTCATGAATTTCTAAAGTAA 8295
DB 1450 GAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAGTAAACATGAAGAAAG 1509
QY 8296 GCTTGATGAAGATTAAGTAACTGTAATCTTGTATGATCAAGAAATGAGAGGA 8355
DB 1510 GAAAAAGTAAACATGAAGAAAGAAAGTAAATCATGAAGAAAGAAAGAAAG 1569
QY 8356 ATTGCCGAGCAGAGTGAATCTTCAAGCAAAAGTAAATCCGATGTTGCT 8415
DB 1570 GATGAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTGA 1629
QY 8416 TTAAGAAACAGATTAATGAATTAATGAATTAACAAAAATATTTCAAGAGTCAAT 8475
DB 1630 GAAGAGAAAGAAAGAAAGAAAGAAAGTGAAGAAAGAAAGAAAGAAATGAGAA 1689
QY 8476 GCTCTGCTTTTAAATGATCAAAAGATGAAGCAATATGAATCTTAAAGGTAGCCGT 8535
DB 1690 GAAGAAAGAAAGAAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1749
QY 8536 GTGCATGCAAGAGAAACAAAGCAATTAAGAGATCAACAAAGTTAACTTTACACT 8595
DB 1750 GTAAACATGAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAGAAAGTAAACAT 1809
QY 8596 GTAAATGAGAGAAAGTATCTCAACTTCGCAAAAGCTTTGCTTAAATGAAATTAAT 8655
DB 1810 GAAGAAAGAAAGAAATTAATGAAGAAAGAAAGAAAGTAAACATGAAGAAAG 1869
QY 8656 GGAATGTAAAGAACTGAGAGAGCTTATGTCGAGCGGAAACAGACCGTTGAAAT 8715
DB 1870 GAAGAAAGTAAACATGAAGAAAGAAAGTAAACATGAAGAAAGAAAGTAAAG 1929

Db	2659	GAAGAAAGTAAACACTGAAAGAAAGAAAGAAAGTAAATACATGAAGAAAGAAAGAAAGAAAGAA	2718
Qy	8356	ATTGCCGAGACGAGAGTGAATATCTTCTACACACACAAAGTAACTAGATCCGATGTT	8415
Db	2719	GATGAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGATGAG	2778
Qy	8416	TTACGAAAGCAAGATTATGAAAAATATATGATTACACAAAAAATATATTTTCAGAAAGTCAAT	8475
Db	2779	GAAGAGGAAGAAAGAAAGAAAGAAAGATGACGAGAGAGAGAGAAAGAAAGAAAGAAATGAGGA	2838
Qy	8476	GCTCTTGCTTTAAATGATATACAAAGATGAAGCGAATATAGAAATCTTTAGCGGTAGCCGGT	8535
Db	2839	GAAGAGAGAAAGAAAGAAATATAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGACATGAAGAGCA	2898
Qy	8536	GTGCATGCACACAGAGAACAAACAAAGCATTTTACGAGATCAACAAAGTTAACTTCTACACT	8595
Db	2899	GTAACACATGAAAGAAAGAAAGAAAGAAAGTAAACATGAAAGAAAGAAAGAAAGTAAACAT	2958
Qy	8596	GTAATATGAGAGAAACGTATCTCAACTTCGTGCAAAAAGCTTTTGCTTAAATATGAAATATAT	8655
Db	2959	GAAGAGAGAAAGAAATATGTAACATATGAAAGAAAGAAAGAAAGAAAGTAAACATGAAGAGCA	3018
Qy	8656	GGAATATGTAAGAAAGAACTGAGAGAGCCTTAGTCGAGCGGAAACAGCAGCCGTTGAAAT	8715
Db	3019	GAAGAAAGTAAACATGAAAGAAAGAAAGAAAGTAAACATGAAAGAGAAAGAAATGTTACCA	3078
Qy	8716	TATACAAAGATTACTACAGAGACATGCTGTCAGAGAAATTTGGAAATTTGAGATTAATTA	8775
Db	3079	TATGAGAGAAAGAAAGAAAGAAAGTAAACATGAAAGAAAGAAAGAAAGTAAAGAAAG	3138
Qy	8776	GAAGAGATGCAAGAGATATATCGATTGTTAAAGATCAACGAGACGG	8822
Db	3139	AAGAAAGATTAATGAT	3185
RESULT 35			
AC122056			
LOCUS	AC122056		
DEFINITION	Mus musculus BAC clone RP24-466A3 from chromosome 14, complete		
ACCESSION	AC122056	179685 bp	DNA linear ROD 13-NOV-2003
VERSION	AC122056.4		
KEYWORDS	HTG.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		
	Sciurognathi; Muridae; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 179685)		
AUTHORS	Kyung, K., Doebber, A., and Kim, K.		
TITLE	The sequence of Mus musculus BAC clone RP24-466A3		
JOURNAL	Unpublished (2001)		
REFERENCE	2 (bases 1 to 179685)		
AUTHORS	Wilson, R.		
TITLE	Sequencing of Mus musculus		
JOURNAL	Unpublished (2001)		
REFERENCE	3 (bases 1 to 179685)		
AUTHORS	McPherson, J.D. and Waterston, R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park		
REFERENCE	4 (bases 1 to 179685)		
AUTHORS	McPherson, J.D. and Waterston, R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park		
REFERENCE	5 (bases 1 to 179685)		
AUTHORS	McPherson, J.D. and Waterston, R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-OCT-2002) Genome Sequencing Center, 4444 Forest Park		
REFERENCE	6 (bases 1 to 179685)		

AUTHORS Wilson, R.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Oct 11, 2002 this sequence version replaced gi:22476311.

Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu

Summary Statistics
Center project name: M_BB0466a03

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION: The RPc1-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone

```

FEATURES
source
Location/Qualifiers
1..179685
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="14"
/map="14"
/clone="RP24-466A3"
/clone_1lb="RPci-24"
1392..1452
/rpt_family="B4"
1489..1616
/rpt_family="ERV.L"
2081..2289
/rpt_family="ERV.K"
2311..3896
/note="Unresolved simple sequence repeat."
3884..4165
/rpt_family="ERV.K"
3919..4254
/note="Sequence derived from one plasmid subclone."
4752..4976
/rpt_family="L2"
7171..7369
/rpt_family="MER2_type"
7418..7585
/rpt_family="MER2_type"
7734..7942
/rpt_family="L1"
7981..8384
/rpt_family="L1"
8584..8891

```

repeat_region	/rpt_family="MaLR"	8892.._9116
repeat_region	/rpt_family="MBR1_type"	9117.._9296
repeat_region	/rpt_family="MaLR"	9581.._9630
repeat_region	/rpt_family="ERV1"	10713.._11357
repeat_region	/rpt_family="L1"	14287.._14580
repeat_region	/rpt_family="MaLR"	14603.._14661
repeat_region	/rpt_family="MaLR"	14785.._14836
repeat_region	/rpt_family="MIR"	16610.._16701
repeat_region	/rpt_family="ERV1"	17755.._18068
repeat_region	/rpt_family="MaLR"	19826.._20270
repeat_region	/rpt_family="L1"	21108.._22054
repeat_region	/rpt_family="RMR13B"	22306.._22358
repeat_region	/rpt_family="Alu"	23799.._23868
repeat_region	/rpt_family="ID"	25018.._25101
repeat_region	/rpt_family="MBR2_type"	25568.._25637
repeat_region	/rpt_family="L1"	25613.._26022
repeat_region	/rpt_family="L1"	27955.._28008
repeat_region	/rpt_family="MIR"	29166.._29220
repeat_region	/rpt_family="L2"	29681.._29735
repeat_region	/rpt_family="ERV1"	29736.._29797
repeat_region	/rpt_family="L1"	32427.._32540
repeat_region	/rpt_family="MBR1_type"	32609.._32960
repeat_region	/rpt_family="L1"	32956.._33003
repeat_region	/rpt_family="ERV1"	33066.._33405
repeat_region	/rpt_family="RMR10A"	34567.._34921
repeat_region	/rpt_family="L1"	35286.._35478
repeat_region	/rpt_family="MIR"	35939.._36745
repeat_region	/rpt_family="CR1"	37053.._37182
repeat_region	/rpt_family="B4"	38641.._38787
repeat_region	/rpt_family="Alu"	39175.._39281
repeat_region	/rpt_family="MaLR"	39322.._39370
repeat_region	/rpt_family="ERV1"	39466.._39728
repeat_region	/rpt_family="B4"	39768.._39914
repeat_region	/rpt_family="Alu"	39966.._40159
repeat_region	/rpt_family="MaLR"	40180.._40391
repeat_region	/rpt_family="ERV1"	40402.._40578
repeat_region	/rpt_family="ERV1"	

[illegible]

Db	3145	GAAG
----	------	--

JACKSON, L., JACOB, L., JIANG, H., JOHNSON, B., JOHNSON, R., JOIVET, A.,
KARPECH, S., KELLY, S., KELLY, S., KHAN, Z., KING, L., KOYER, C.,
KOWIE, C., KRAFT, C.L., LEDOW, H., LEVAN, J., LEWIS, L., LI, Z., LIU, J.,
LIU, J., LIU, W., LIU, Y., LONDON, P., LONGACRE, S., LOPEZ, J.,
LORENSAUER, L., LOULESSED, H., LOZADO, R.J., LU, X., MA, J.,
MAHEBHARI, M., MAHINDARTNE, K., MAHMOUD, M., MALLORY, K., MANGUM, A.,
MANGUM, B., MAPUA, P., MARTIN, K., MARTIN, R., MARTINEZ, E.,
MAWHNEY, S., MCLEOD, M.P., MCKEILL, T.Z., MEENAN, E.,
MILGOALJEVIC, A., MINER, G., MINJA, E., MONTENAYOR, J., MOORE, S.,
MORGAN, M., MORRIS, K., MORRIS, S., MUNDASA, M., MURPHY, M., NAIR, L.,
NANKERVIS, C., NEAL, D., NEWTON, N., NGUYEN, N., NORRIS, S.,
NWAOKOLEMEH, O., OKUNO, G., OLANPUNGBON, A., PAL, S., PARKS, K.,
PASTERNAK, S., PAUL, H., PEREZ, A., PEREZ, L., PFANKOCH, C.,
PIOPPER, P., POINDESTER, A., POPOVIC, D., PRIMUS, E., PU, L., -L.,
PLAZO, M., QUIROZ, J., RACHLIN, E., REEVES, K., REGIER, M.A., REIGH, R.,
REILLY, B., REILLY, M., REN, Y., REUTER, M., RICHARDS, S., RIGGS, F.,
RIVER, C., RODKEY, T., ROJAS, A., ROSE, M., ROSE, R., RUIZ, S.J.,
SANDERS, W., SAVERY, G., SCHERER, S., SCOTT, G., SHATMAN, S., SHEN, H.,
SHETTY, J., SHVARTSBEYN, A., SISON, I., SITTER, C.D., SMALIS, D.,
SNEED, A., SODERGREN, E., SONG, X.-Z., SORELLI, R., SOSA, J.,
STEINLE, M., STRONG, R., SUTTON, A., SVATEK, A., TABOR, P., TAYLOR, C.,
TAYLOR, T., THOMAS, N., THOMAS, S., TINGEY, A., TRELOS, Z., USMANI, K.,
VALAS, R., VERA, V., VILLASANA, D., WALDRON, L., WALKER, B., WANG, J.,
WANG, Q., WANG, S., WARREN, J., WARREN, R., WEI, X., WILEY, F.,
WILLIAMS, G., WILSON, R., WLECZYK, R., WOODEN, H., WORTH, K.,
WRIGHT, D., WRIGHT, R., WU, J., YAKUB, S., YEN, J., YOON, L., YOON, V.,
YU, F., ZHANG, J., ZHOU, J., ZHOU, X., ZHAO, S., DUNN, D., von
NIEDEHANSSEN, A., WEISER, R., SMITH, D.R., HOLT, R.A., SMITH, H.O.,
WEINSTOCK, G. and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 233713)
Unpublished
Direct Submission
2 (bases 1 to 233713)

JOURNAL
TITLE
AUTHORS
REFERENCE
JOURNAL
TITLE
JOURNAL
REFERENCE

Submitted (06-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 233713)

	RESULT	36				
	ACI33034					
	LOCUS					
	DEFINITION	Rattus norvegicus clone CH230-6IH4,	DNA	linear	HTG 19-NOV-2002	
	ACCESSION	ACI33034				
	VERSION	ACI33034.2				
	KEYWORDS	HTG; HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.				
	SOURCE	Rattus norvegicus (Norway rat)				
	ORGANISM	Rattus norvegicus				
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Myrodonta; Muridae; Murinae; Rattus. 1 (bases 1 to 233713) Muzny D.Marie, Metzker M.Lee., Abramson S., Adams C., Alder J., Allen C., Allen H., Alperbrooks S., Amin A., Angiano D., Anyalectchi V., Ayogagi A., Ayodeji M., Baca E., Baden H., Bailwin D., Bandaranaike D., Barber W., Barnstead M., Benhammed F., Bliswal K., Blair J., Blankenburg K., Blyth P., Brown M., Bryant N.K., Bunyah C., Burch P., Butrell K., Calderon E., Cardenas V., Carter K., Cavazo I., Ceasar H., Center A., Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J., Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., De Souza L., Davila M.L., Davis C., Day Carroll L., De Anda C., Dedetich D., Delgado O., Denison S., Deramo C., Ding X., Dim H., Divya K., Diaper H., Dugan Rocha S., Dunn A., Durbin K., Duval B., Evans K., Egan A., Escotto M., Eugene C.A., Evans C.A., Falls T., Fan G., Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P., Frazer C.M., Gabisi A., Ganter R., Garcia A., Garner T., Garza M., Gebeorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W., Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K., Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez U., Hernandez R., Hines S., Hlaung S.L., Hodgson A., Hogues W., Hollins B., Howells S., Huylk S., Hunne J., Idlebit D., Jackson A.				

JOURNAL
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 COMMENT
 On Nov 19, 2002 this sequence version replaced g1:22748196.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/atl/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KAZP
 Center clone name: CH230-61H4
 ----- Summary Statistics
 Assembly Program: Phrap, version 0.990329
 Consensus quality: 217928 bases at least Q40
 Consensus quality: 220670 bases at least Q30
 Consensus quality: 222620 bases at least Q20
 Estimated insert size: 226602; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).

- * NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

1 233713: contig of 233713 bp in length.

FEATURES

source
1..233713
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-61H4"
1..1147
/note="wgs_end_extension
clone_end:T7"
19..750
/note="clone_boundary
clone_end:T7"
site:
end_sequence:BH283618"
complement(231046..231754)
/note="clone boundary
clone_end:Sp6
site:
end_sequence:BH283619"
232397..233713
/note="wgs_end_extension
clone_end:Sp6"

misc_feature
1.0%; Score 101.6; DB 14; Length 233713;
Beef Local Similarity 36.1%; Pred. No. 9.8e-05;
Matches 454; Conservative 0; Mismatches 803; Indels 1; Gaps 1;

ORIGIN

Query Match 1.0%; Score 101.6; DB 14; Length 233713;
Beef Local Similarity 36.1%; Pred. No. 9.8e-05;
Matches 454; Conservative 0; Mismatches 803; Indels 1; Gaps 1;

7905 AGAATAAGATGATCTAGCCCTTCTTAAAGAAATTGGAAGAAATCAATGTCAAGTGA 7964
202313 AAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 202372

7965 AAAAGAAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8024
202373 AGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCA 202432

8025 AGGAATTATTTCCGAAGCAAAAGATGCCGAAAGCTCTTAATTGAAAATTGATCAAAATC 8084
202433 AGAAGAAAGAAAGAAAGAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 202492

8085 CGAAGAAAGTATTTTCATGCAGATTAATGTGAATATGAAAGCAACATTAATAAGT 8144
202493 AGAAGAAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGC 202552

8145 AACAGCAGTTTCTAAAGCAGTAACAGTTCTGTATTGGAGAGATTGAGTCAACCAAGC 8204
202553 AGAAGAAAGCAAGAAAGCAAGAAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCA 202612

8205 AGAAGCTACTGCTGCAAGTAAATCTATGTGAAGTTGAGAAAGAAATTTGTTCAAGAC 8264
202613 AGAAGCAAGAAAGAAAGAAAGAAAGAACACAGAAAGAAAGAAAGAAAGCAAGAAAG 202672

8265 AAATGATTTGAATGCAATTTCTTAAGTAAAGTTTGAATGAATTAAGTAACTGCTTA 8324
202673 AGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAG 202732

8325 ATCTTCTGTAGTATCAAGAAATGAGAGAAATTTGCCGAGAGCAAGAGTGAATCTTCTAC 8384
202733 AGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAG 202792

8385 AGCACAAGTAATATCTGAATCCGTAGTTGTTTACAAAACAAAGTTATGAATAATGA 8444

Db 202793 AGAAGAAAGCAAGCAAGAAAGAAAGAAAGCAAGCAAGAAAGAAAGAAAGAAAGAAAG 202852
Oy 8445 TTACCAAAAAATATATTTTCAGAGTCAATGCTCTTCTTTAATGATACAAAGATGA 8504
Db 202853 AGAAGAAAGAAAGAAAGAAAGAAAGCAAGAAAGCAAGAAAGAAAGAAAGAAAGC 202912
Oy 8505 AGCGAATATAGAAATCTTTAGCCGTAGCCGATGTCATGCACAAAGAAACAAAGCAAT 8564
Db 202913 AGAAGAAAGCAAGCAAGAAAGCAAGAAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAG 202972
Oy 8565 TACGAGATCAACAAAGTTAATCTTCAACACTGTAATGAGAAACGTAATCTCAACTTCG 8624
Db 202973 AGAAGAAAGCAAGAAAGAAAGCAAGCAAGAAAGAAAGAAAGAAAGCAAGAAAGC 203032
Oy 8625 TGCAAAAGCTTTGGCTAAATTAATTAATGAAATGTAATGAAAGAACTGAGAGACCTT 8684
Db 203033 AGAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGAAAGAAAGAAAGCAAGAAAGCAAG 203092
Oy 8685 AGTCGAGCGGAAACAGCAGCCGTTGAAATTTATCAAAAGATCTACAGAGCAATTGCT 8744
Db 203093 AGAAGAAAGAAAGAAAGAAAGCAAGAAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAG 203152
Oy 8745 TGCAAGAAATTTGGGAAATTTGAGATTAATTTAGAAACGATTGCAAGATTAATCAATTGT 8804
Db 203153 AGAAGAAAGAAAGAAAGAAAGAAAGCAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAG 203212
Oy 8805 AAGAGTCAACGAGACGGAACCAAGAGAGTCTGTGCGAAAGAAATGATTTCTGTGA 8864
Db 203213 NNN 203272
Oy 8865 AAATACAAATTTCAAGGAGAAACAAATTCATCAATGAAGATAAAGCCAGAAATTGTTGAAC 8924
Db 203273 NNN 203332
Oy 8925 CGAAGGTAAATATGATGCTTTGAATGAACCTTGATGATGATCAAGAAAGAAAGTGG 8984
Db 203333 NNN 203392
Oy 8985 TGCGTATGTGGAATTTGATATTTGAAATGTTGATGTAATATGATTAAGAAATATGT 9044
Db 203393 NNN 203451
Oy 9045 AGAAGCCAAATTCGAGAGACATGCTATTGTAGAAACTACTGAAAAACAAGAAATATCAAGC 9104
Db 203452 AGAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGAAAGAAAGAAAGAAAG 203511
Oy 9105 ATTTCAGAGCAAAAGTAATATTTCTTGGAAGAAAGAAAGCAAGCTGCAAGCTGCA 9162
Db 203512 AGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 203569

RESULT 37

LOCUS BX890609 228713 bp DNA linear VRT 29-APR-2004
DEFINITION Zebrafish DNA sequence from clone DKEY-56E5, complete sequence.

ACCESSION BX890609
VERSION BX890609.7 GI:46878891
KEYWORDS HTG

SOURCE Dario reio (zebrafish)
ORGANISM Dario reio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
Cypriniformes; Cyprinidae; Dario.

REFERENCE 1 (bases 1 to 228713)
Health, P.

AUTHORS Direct Submission
TITLE Submitted (29-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 29, 2004 this sequence version replaced gi:46275351.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC

Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of monoclonal A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhiting Bao and Sean Eddy, submitted), and those beginning 'drv' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_terio/fishmark.shml DKEX-5055 is from a zebrafish BAC library

VECTOR: pindigBAC-5.

FEATURES	Location/Qualifiers
source	1. .228713

ORIGIN

Query Match	1.0%;	Score 101.4;	DB 5;	Length 228713;
Best Local Similarity	43.6%;	Pred. No. 0.00011;		
Matches 938: Conservative	0;	Mismatches 1191;	Indels 22;	Gaps 107;

Qy	7110	GGAAAAGAAATGTATACGGCAAAAGAGAAAAAATATTAAGCAACACGACCAATCG	7169
Db	152851	AGAGGTTGAAGAAAGAAAGAAAGATGATATGATGTGGTGTGAAAGAAAGAAATAAAG	152792
Qy	7170	AGGAATCGAGAGACGCAATCGAGCCAAATCTTGGTAAATTAATTTTGAAACAGCTGT	7229
Db	152791	TGAAGAGGTTGAAGAAAGAAAGAGATGATATGCAAGAAAAGAAATTAAGATGAATGA	152732
Qy	7230	AGAAATGAAAAAAATTTCTGAAGAAAAAGAAACAGAACTTTTAAACCTTTAGACGAAGT	7289
Db	152731	AGAGGTTGAAGAAAAATATGATGATGATGATGACAGAAAAAGAAATAAAAAGAAATAAAG	152672
Qy	7290	TAAACAAGAACAAAGTAAAAAAAGTAATGATCTACGAAAAAAATCTTCAATCAGACAG	7349
Db	152671	GGTTGAAGAAAGAAAGATGATGATGATATGCAAAAAAGAAAAAGATGAAAAATTTGA	152612
Qy	7350	TATTTCTACAGAAAGATCTTCTGTAAACCGGATAGAGAGATCACTCAGGAGAAAGAAAT	7409
Db	152611	TGAAGATGATGAAGAAAGATGCAATAATTAGAAAGTGAATGCAGAAAAAAAAGA-----GAT	152558
Qy	7410	TAAAGCAATGTGTAAGACTTCTGATATTAATTGAAAAAAATGTAGATTTACACAGAGGA	7469
Db	152557	GAAAGATGATGCAATGAAAAATGAAAGAACAGATTTGAAGATGGAAGATGCAAGTTGATGAAGATGA	152498
Qy	7470	CAGAATAATATCACTTCTAATCTGTGGTTGGGAATCGCAGAGCTTGGCTTCGGCATCAG	7529
Db	152497	TGATGAAAAAGAAAGATTAAGAAAGTCAATGAAAGAGATGAAGATATAGTTGATGCAAAA	152438

QY	7530	AAACGTGCACTTCAAAATATTTAAAGAAATCCGAGCTTACTGTGTGAAATTTCTTTGT	7589
Db	152437	AAGA----AATTAAGTTGATGAGAGAGATGATCGAGAAAAGAAATAGATGAAGATGA	152383
QY	7590	GAAAGCAAGCTGAAAAAGTAAATGTTAGTCCGATATTACAGAAATGTTCCTTAAACGC	7649
Db	152382	AAAGTTGATGAAGATGAAAGATGTGGAAGAGA--AGAAAAAAACATGAAATCATGTAGT	152324
QY	7650	ATATCAAGTCTCTGTAGAGCAATTGGGAATAGAGTGCCTTATCGAAGATTTAAATTTCTAA	7709
Db	152323	GGATGCGAATAATGAAGATGATGCGAAAAGGAGCAGATGAAGTATGAGAGATGATGA	152264
QY	7710	TGGAAGATCAATATTCAGTATTTAAAAATTTCTAAGCTATTAGAAAAAATATATGATTAT	7769
Db	152283	TGAAAAAGACAAATTTGATGGAAGAGAAATATGATCGAABAAAAGAAATGAAGGAAGATTA	152204
QY	7770	TGTAAAGATTAATTCGAAATTGAGACGGAAACAAAGATTAAACCGTATGAGCCGTAGC	7829
Db	152203	GGTTGATGAAGAAATCAAGTTTAAAGAAAGATGCGAAAAGATPAAAAAATGAAAGCGTTCA	152144
QY	7830	TGCCGAGCCATTATCTCTAAAAAGCAAAATGAAATGAAATTCAGAGGTTGAAATTGAGAA	7889
Db	152143	TGAACACATTAAGAACAAAGATGCGAAATTTAAAGATTAATGCGA--AAAGTAGCAGATGA	152085
QY	7890	GAGTATTTTCAATGAGAAAAATAGAGTAACTAGCCCTTTAAAGAAATTGAAAGAGAAAT	7949
Db	152084	AGCTGATGATATATGATGAAGAAATTGATGAAGAACAGATGATGAGATGAAAGATTAA	152025
QY	7950	CAATGTCAAATGTGAAAAAGAAACAGAGTCACTGCTGATTTCAAAGAGCTTCTGTAGG	8009
Db	152024	CAAAAGTATATGATGATGATGCAAAAA---AGTACCAAACGAAAGAAAGATGAGAAAA	151969
QY	8010	AGCAGTAGACAGGCGCAGGAATTATTTCCGAGCAAAAGATGCCGAGCTCTTATTTGAA	8069
Db	151968	AGAACAAAGTGAAGAACAGAGATGATGAABAAAAGAAATATGATGCTGAAAAGAAAGAAATAG	151909
QY	8070	AGTTAGTACAAAATCCGAGAGAGTATTTTTCATCGAGTAAATGGAATATGGAAGCAAC	8129
Db	151908	AAGGTGAAGAGATGATGCGAACAAGTAAATGAAGAGAGATGAAGAAAGTTGATTAAGAAAG	151849
QY	8130	ACATTAATGAAAGTAAACAGCACTTTCTAAACACATTAACAGTTCTGATTGGGAGAGT	8189
Db	151848	AAATTAAGAT--CATGACAGAGAAAGAAAGAAAGAAATGAAGATGGAAGATGAGACAAAG	151790
QY	8190	TGAGATCACCAAGCAGAAAGTACTGCTGCAGGTAAACCTATGTGTGAAGTTGAGAAAG	8249
Db	151789	TCATGCAAGCAAAAGAAAAGATGAAAGAGTTGATGAAGGAAAGAAATTTGAAAGTGAAG	151730
QY	8250	AAATTTGTTCAGAACAAATCGATTGATGCAATTTTCTAAAGTGAAGGTTTGGATGAAG	8309
Db	151729	AATGATGAAGAAAGATGATGATATGAAGATTAAGT--TGATCGAAGAAAAGAAATTAAGAG	151671
QY	8310	TAAAGTAATCGTAAATCTTCTGTGTATCAGAAATGAGAGAGAAATTGCCGAGCAGG	8369
Db	151670	AGATGAGTTGATGAAGAAAGATGAAGTTAAAGAAATGCGAAGAAAGTAAAAATGAAGA	151611
QY	8370	AGTGAATATCTTCTTACAGCAAAAGTAAATCTGAATCCGTAGTTGTTTTCGAAGCAAGA	8429
Db	151610	CGTTGATGAAGACGATTAAGAAAGAAAGATCGAABAAATGAAGCAATTAAGCAGAAAAGTAGC	151551
QY	8430	TTATGAAATTAATGATTAACAAAAAATATATTTTCAGAAAGTCAAGCTTCTGTTTAA	8489
Db	151550	AGATGAAGGTGATGATTAATGATGAAGAAATTGATGAAGAGCAGATGATGAGATGTTAA	151491
QY	8490	TGATCAAAAGAAATGAAGCAATATAGAACTTTAGCCGTAGCCGCTGTGCAATGCAACAG	8549
Db	151490	CGTTTATGAAGGTGATGATGATGATGAAGAAATTAAGTTGATCAAGAAACAAAGATGAAGT	151431
QY	8550	AACAAACAAGCAATTTACAGAGATCAAAACAAGTTAATCTTCAACAATCTGTAATCGAGAAA	8609
Db	151430	TGAAGTGAAGAAAGATGATCGAAGAAAGAAATTAAGCAATGAAGAAAGATCGAAGAAATGA	151371

QY	8610	GGIATCTCAACTTCGCGCAAAAAGCTTGGCTAAATAATGAAATATATGAAATGCTAAAGG	8669
Db	151370	AGATATGCGAAGAAAGAGCGCGATAGAAAGATATGATGAGCGTTGTGAATPAGAAAG	15131111
QY	8670	AACTGAGAGAGCCTTAGTCGAGCCGAGAAACAGCAGCCGTGGAAATTTATACAAAGATAC	8729
Db	151310	TGATCGAAGAAAGAAAAGATGAGATGAGAGAGTTGAAGAGATTAAGAGAGATGAG-	15125252
QY	8730	TACAGAGACATTGGTTCGAGAAATTTGGGAATTTGAGATTAATTGAAACGATTGCAAG	8789
Db	151251	TGCAGAAAAAAGAACAGAGATGTTGTTGATGATAGACGATGAAAAAATAGGGAATGGAAG	15119292
QY	8750	AGATATATCGATTTGTAAAGTCAACCGAGACCGAACCAAGGAGGTCTTGCGG-AAAG	8848
Db	151131	ATGATGCAAAATAGTATGCGAGAGAGAGAGAGAGATGATGATGATGATGATGATGAAAG	15113222
QY	8849	ATGCTATTCTCTGTGAAAAAATACATTTTCAGGGGAAACAAATATCATTCATTGAGATTAAG	8908
Db	151131	ACAAGATGATGAGAAAGATATGATCTGTAAGAGAGAGATATGAGATGAGAGAGTTG	15107272
QY	8909	CCAGATTGTTGGAACCGAGATGTAATATGATGCTTTGAATGAACTTGATGTAGATC	8968
Db	151071	ATGAAGAGATGTAGATGAAAAACAATATGATTAGTAAAGAAAGAAAGAGAGAAAT	15101212
QY	8969	TACAAGAAAAAGGTGGCTATGGTGGAAATTTGTTGGAAATGTTGATGTAATTAAG	9028
Db	151011	ATGATGAGAAAAAGAGATTAAGAT-GAAGATGATGTTGAAAAAATAATGAAAGAGATGAAG	15095333
QY	9029	TGATTAGAAAAATGTAGAACGCCAAATTCGAGAGCATCTATTGTAGAAACTACTGAA	9088
Db	150952	TGGATGATTAAGAGATGACGTTAAAAAAGATGCAAAAGAGTAAAGATGAAGACGTTG	15089333
QY	9089	AACAAGATATTAAGCATTACAGAGCAAAAGTAATTTCTTGGAAAAAGAGCGCTG	9148
Db	150892	ATAAAGACGATGAAGAAACAGATGCGAGAAAAATTAAGATGAAGAGATTAAGATGAAA	15083333
QY	9149	CAGCTGCAGCTCCAAATATGCAATGTCACATTTCCAAATGAGATGATATTAATAATTGG	9208
Db	150832	TAGGTGGGATTAAGAACTGATGAAAGATTTATGACAAAGAGATCTAAAGAAAAGAG	15077333
QY	9209	CAAGACGATGATCTTCTCAATTATTAACCAAAATTTCAAAAATTAATTA	9259
Db	150772	AAAAATTAAGATGATGATGATTAATGATGCTGAAGAAAGAGAAAGAA	150722
RESULT 38			
AF270648			
LOCUS	AF270648	4102 bp	DNA linear INV 12-JUN-2000
DEFINITION	Plasmodium falciparum mature parasite-infected erythrocyte surface		
VERSION	AF270648		
KEYWORDS	AF270648.1 GI:8468620		
SOURCE			
ORGANISM	Plasmodium falciparum (malaria parasite P. falciparum)		
REFERENCE	Plasmodium falciparum		
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
TITLE	1 (bases 1 to 4102)		
FEATURES	Shen,Z.X., Yu,X.B., Fang,J.M., Li,X.R., Ma,C.L. and Wu,Z.D.		
JOURNAL	Cloning and expression of the mature parasite-infected erythrocyte		
REFERENCE	surface antigen gene of Plasmodium falciparum isolate FCCI/HN		
AUTHORS	Unpublished		
REFERENCE	2 (bases 1 to 4102)		
JOURNAL	Shen,Z.X.		
TITLE	Direct Submission		
FEATURES	Submitted (20-MAY-2000) Department of Parasitology, Sun Yat-Sen		
REFERENCE	University of Medical Sciences, 74 Zhongshan Road II, Guangzhou,		
FEATURES	Guangdong 510089, P.R.China		
REFERENCE	Location/Qualifiers		
FEATURES	1..4102		
REFERENCE	/organism="Plasmodium falciparum"		
REFERENCE	/mol_type="genomic DNA"		
REFERENCE	/isolate="FCCI/HN"		
REFERENCE	/db_xref="taxon:5833"		

gene	CD5	mRNA
<p> county="China" <1..>4102 /gene="MESA" join(<1..150,281..>4102) /gene="MESA" /product="mature parasite-infected erythrocyte surface antigen" join(1..150,281..4102) /gene="MESA" /codon_start=1 /product="mature parasite-infected erythrocyte surface antigen" /protein_id="Aa75554.1" /db_xref="GI:8468621" /translation="MEYICRNLCYDKNNMENEKVKVYVNNSLKYMKFLCTTI ITFCLDIDYNCSPYYSISIKNNDRVYRILISETPMSLSERKTPEDHLYSIR NYIECLNAPYIDDPWGVSVDKRRNCCLOHILILEVQASERKQOESNADISEI RKEKEKYMKEIEMDSRDYKAFRLQFLKIDRPRPHVMMBSKRTDSSKRTD ESKRTDSKRTGSKSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKET TGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETG LEKRYTKITSEPFETERVKITTESKORBGKVGSPYENSNVSSESSETKVLAKS ENGEKRGKGVNVDGASNSBDDPKKLPEOENGTGYESSESESESESESESESE KKKKSPFOMLGCNPLCNKIETDDESETLVVDNKRKKFLRANTEKNDKSKK LDRGKEDYKENDDKDKVIGDDESEKVDKNDOKVIGEGKEDYKENDDK DKYIGSEKTKQKEIKERYKVRKCKKRYKKGIKENDTBENDYKRPETIIEVKSRI IKQVDEIKENDTBENDYKRPETIIEVKSRIIEKQVDEIKENDTBENDYKQGOI1 TEBEKEIKQVDEIKENDTBESKDKLIGQSI1IEBEVKIKENDTBENDYKQGOI1 KVIQGOI1IEBEVEKEIEKQVDEIKENDTBESKDKVIGQEV1KQVDEIKENDTB OEVRKYKKEKKEKKEKVKKRNKNKNKNVIGKEMEDVNEKDTANDKY1BEOK EKSEVKEVKEKDTVNDKVIQGOV1IEBEVKEKVKVKNVKNRKNRKNRKNVIGETV DVEKDTESKDKM1IGKEV1IEBEVEKRVKNVKNRKNRKNRKNRKNVIGETV EAVNEKDTKNNDKK1IGRVKRP1DDCKKEKEVESESESESESESESESESESE ESE ESE SKHNTLNTFRPMNALNEDPREFHNILSIHEDTLMELKRIILVNLPLENPHNNOK AEIDKFRSEMNVVHQLNTEIRIMYBENARBLNTVILDPITITNVIFGDDTMKFI IDE" </p>	<p> 1.0%; Score 100.4; DB 2; Length 4102; Best Local Similarity 43.9%; Pred. No. 0.0064; Matches 572; Conservative 0; Mismatches 721; Indels 9; Gaps 3; </p>	<p> ORIGIN </p>

[illegible]

Journal	Patent: EP 1526178-A 11 27-APR-2005;
Features	Institut Pasteur (PR)
Source	Location/Qualifiers
	1. .1701
	/organism="Plasmodium falciparum"
	/mol_type="unassigned DNA"
	/db_xref="taxon:5833"
misc_feature	1. .1701
	/note="MSP 3.6"
Query Match	1.0%; Score 98.6; DB 6; Length 1701;
Best Local Similarity	42.8%; Pred. No. 0.0015;
Matches	610; Conservative 0; Mismatches 809; Indels 6; Gaps 2;
Db	7052 CTATAAATAGTCTGTGATAGCTGAGTTCACATTAATCTGTATTGACTTCGGTACGAG 7111
Db	110 CTTTAAACGAAATTTAAGAAATGAGATTTAAATATATACCTTATTTTAAAGAAACAA 169
Qy	7112 AAAAGTAAATGTAACGCGAAAAGGAAAAAATATTAAGCAACAGCAGCAATGACG 7171
Db	170 ATATATTTAAATATAGAAATGAAATTAATATCAATTTATATGAGTACAGAGAA 229
Qy	7172 GAATCGAGAGGAGCAGCAATCGAGCCAAATGCTTGTGTAATTAATTTTGAACAGCTGAG 7231
Db	230 CTAAAGAAATTTATGATATTAATGAAAAATTTTCCGTGATATTTTCTTGATTA 289
Qy	7232 AAGATAGAAAAATTTGGAAGAAAAGAAACAGAGTTTAAATACTTTAGCAGAGTTA 7291
Db	290 TCTTTACGAAAAATAAAGAACAAAAAATGAGAAAGTACCAATGAAATGAAATGATAA 349
Qy	7292 ACAAGAAACAAGATTAATAATGATCTACGAAAAAATCTTCAATCAGCAGATA 7351
Db	350 ATGATGAGAAAGATTAACAAATATGATTTGAAAAAATGAGAAAGTGAATAA 409
Qy	7352 TTTCTACGAAAGATTAATCTGTAAAGCGATAGAGAGATCTCAGGAGAAAGAAATTA 7411
Db	410 AATCGAAACTGAATATAGTGAAGAAATTAATCTGAAAAAGTATGAAAAAGTACCTGAAG 469
Qy	7412 AAGCCATTGTGAGACTTGTGATATTTATGAAAAAATGTGATATTAACAACAGAGCA 7471
Db	470 AAGTAGCTGAAGAAATTAAGTGAAGAAAGTAGAGAAAGTACGAAAGATTAAGTTGAA 529
Qy	7472 AGAATATATCACTTCACTGATGCTTGGGAACTGAGAGCTCTGCTCCGCAACAGGAA 7531
Db	530 AAGTAGATGAAAAAGTAGCTGAAGAGATCAAAAAGTAGATGAAGAAATTAAGTGAAG 589
Qy	7532 CAGTGCAGTATACAAATATTAAGAAATTCGAGATTAATCTGTGAAAAATTCCTTTGGA 7591
Db	580 AATTAATGAAAAAGTAGATGAAGAAAGTAACTGAAGAAATTAATTTGAAAAAGTAGATGAAG 649
Qy	7592 AAGCAGCTGAAAAAGTAAATGTTAGATCGATATTAACAGAAATGTTCTTTAACAGCAT 7651
Db	650 AAGTTGCTGAAGAAATTAATGAAAAAGTAGATGAAGAAAGTTCGAGAAATTAATTTGAA 709
Qy	7652 ATCAAGGTCCTG---TAGAGACTTTGGAAATAGAGAGCTGCTATGAGAAATTAATCTA 7708
Db	710 AGGTAGCTGATGAATTAATTTAAAAAGTAGAGAAAGTTCGAGAAATTAATTTGAA 769
Qy	7709 ATGGAAGATCAAAATATCAGTATTAATAATTTCTAAGCTTATAGSAAAAAATATTGATGTTA 7768
Db	7770 AGGTAGCTGATGAATTAATTTAAAAAGTAGAGAAAGTTCGAGAAATTAATTTGAA 829
Qy	7769 TTGTAAAAAGTAAATCGAATTTGAGAGCGAGCAAAAGATTAACCTGAGAGCGGTAG 7828
Db	830 AAGTAGCTGAAGAAATTAATTTAAAAAGTAGAGAAAGTTCGAGAAATTAATTTGAA 889
Qy	7828 CTGCCGAGCCCTTATCTCAAAAGCAAGATGAATGAATTCAGAGGTTGAAATTTGGA 7888
Db	890 AAGTAGATGAAGAAAGTAACTGAAGAAATTAATTTGAAAAAGTAGATGAAGAAATTAAGTGAAG 949
Qy	7889 AAGATATTTTCAATAGAAAAATAGAGTAACTAGCCCTTTCTTAAAGAAATTTGAAAGAA 7948

Db 950 AATTAAATGAAAAGTAGATGAAAGAGTTGCTGAAGAAATTAATGAAAAAGTAGATGAAG 1009
Qy 7949 TCAATGTCAAAGTGGAAAAAGAAAAAGAGAGTACTGCTGAATCTCAAGAGCTTCTGTAG 8008
Db 1010 AAGTGTCTGAAGAAATTAATGAAAAAGTAGCTGATGAAATTAAGTAAAAAGTAGCTGAAG 1069
Qy 8009 GAGCAGTAGCAGGCGCAGGAATTAATTTCCGAAGCAAAAGATCCCGAAGCTTTATTGGA 8068
Db 1070 AATTAGTGAAGAAAGTAGATGAACAAGTAGCTGAAGAAATTAAGTAAAAAGTAGATGAAC 1129
Qy 8069 AAGTAGTCAAAAATCCGGAAGAAAGATTTTTCATGCAATATATGTGAATATGAAAGCA 8128
Db 1130 AAGTGTCTGAAGAAATTAAGTAAAAAGTAGATGAACAAGTAGTGAAGAGTAGCTGAAG 1189
Qy 8129 CACATTAATGAAAGTAAACAGCAGTTTCTAAAGCAGTAAACAGTCTGTATTTGGAGAG 8188
Db 1190 AAGTAGTGAAGAAAGTAGATGAAGAAAGTAGAAAAAGTACTGAAGAAAGTAGCTGAAGAG 1249
Qy 8189 TTGAGATCACCAAGCAGAGAGCTACTGCTGCAAGTAAAACTATGTGTAAGTTGAGAGAG 8248
Db 1250 TAGCTGAAGAAAGTAGCTGAAGAAAGTAGCTGAAGAAAGTAGCTGAAGAAATTAAGTAAAAAG 1309
Qy 8249 GAAATTTGTTCAGAAC--AAATCGATGATGCAATTTCTAAAGTAGAAGGTTTGATG 8305
Db 1310 TAGATGAAGAAAGTAGCTGAAGAAAGTAGATGAAGAAAGTAGAAAAAGTAGCTGAAGAG 1369
Qy 8306 AAGATTAAGTAACTCTCTAAATCTTCTGTAGTATCAGAAATGAGAGAGGAAATTCGCGAG 8365
Db 1370 TAGTGAAGAAAGTAGATGAAGAAAGTAGCTGAAGAAAGTAGATGAAGAAAGTAGAAAAAG 1429
Qy 8366 CAGGAGTAGATCTTCTACAGCAACAAGTAACTGAATCCGTAGTTCGTTTACGAAAGC 8425
Db 1430 TACTTGAAGAAAGTAGATGAAGAAAGTAGTGAAGAAAGTAGCCGAAGAGTAGCTGAAGAAAG 1489
Qy 8426 AAGATTATGAAATTAATGATTAACAAAAAATATATTTTCAGAAAG 8470
Db 1490 TAGTGAAGAAAGTAGTGAAGAAAGTAGTGAAGAAAGTAGTGAAGAAATTAAGT 1534

RESULT 40
CS082924
LOCUS CS082924 1701 bp DNA linear PAT 18-MAY-2005
DEFINITION Sequence 11 from Patent WO2005040203.
ACCESSION CS082924
VERSION CS082924.1 GI:66349524
KEYWORDS
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1
AUTHORS Druilhe, P., Singh, S., see See, D. and Mejia, P.
TITLE Msp-3-like family of genes
JOURNAL Patent: WO 2005/040203-A 11 06-MAY-2005;
Infect Dis J (FR)
FEATURES
source
1. .1701
Location/Qualifiers
/organism="Plasmodium falciparum"
/mol_type="unassigned DNA"
/db_xref="taxon:5833"
1. .1701
/note="MSP 3.6"

ORIGIN
Query Match 1.0%; Score 98.6; DB 6; Length 1701;
Best Local Similarity 42.8%; Pred. No. 0.0015;
Matches 610; Conservative 0; Mismatches 809; Indels 6; Gaps 2;

Qy 7052 CTATTAATAGTCTGTGATAGCTCGAGTTCAACATTAATCTGATTTGACTTCGTACGAG 7111
Db 110 CTTTAAAAAGCAATTTAAGAAATGATATTTAAATTAATCTTATTTTAAATGAAAGAAACA 169
Qy 7112 AAAAGTAAATGTAACGCAAGAAAGGAAAAAATATTTAAGCAACAGCAGCAATGCGAG 7171

Db 170 ATAAATTAATATAGAAAAAGAAATTAATTAATACAAATTAATATGAAAGTAAACAGAGAAA 229
Qy 7172 GAATCGAGAGCAGCAATCGAGAGCAATGCTTGTGTAATTAATTTGGAAACGCTGAG 7231
Db 230 CTAAAGAAAGAAATTAATGATATTAATGAAATATTTTCCCTGATTAATTTTCTTGATA 289
Qy 7232 AAGATGAAAAAATTTCTGAGAAAAAGAAACAGAAATTTTAAAACTTTAGACGAAGTTA 7291
Db 290 TCTTTACTGAAAAATTAAGAACAAAAAATGAAAGAGTACCAATGAAATAGAAAGTAGTA 349
Qy 7292 ACAAAACAACAGATTAATAAGTAAATGATCTACGAAAAAATTTTACATCAGCAGGTA 7351
Db 350 ATGATGAGAAAGAAAGTAAAAACAGAAATATGATCTGAAAAAATGAGGAAGTAGAAAAATA 409
Qy 7352 TTTCTACAGAAATCTCTGTAAAGGGAATGAGAGATCTCAGGAGAGAGAAATTA 7411
Db 410 AATCGAAACTGAATTAAGTGAAGATTAATCTGAAAAAGTAGATGAAAAAGTAACTGTAAG 469
Qy 7412 AAGCCATTGTGAAGACTTCTGATATTAATTTGAAAAAATGATGATTAACAACAGAGACA 7471
Db 470 AAGTAGCTGAAGAAATTAAGTAAAAAGTAGATGAAGAAAGTAGCTGAAGAAATTAAGTTGAAA 529
Qy 7472 AGAATATATCACTTCTACTGTGTTGGCAATGCAAGTCTTCTCCGATCAGGA 7531
Db 530 AAGTAGAGAAAAAGTAGCTGAAGAAAGTAGATCAAAAAAGTAGATGAAGAAAGTAACTGTAAG 589
Qy 7532 CAGTGGCAGTTACAAATTTTAAAGAAATTCGAGAGTTACTGTGAAAAATCTTTTGGA 7591
Db 590 AATTAAATGAAAAAGTAGATGAAGAGTAGTACGAAAGATTAATTTAAAAAGTAGATGAAG 649
Qy 7592 AAGCAGCTGAAGAAAGTAAATGTTAGATCGATATTAACAGAAATGTTCTTTTACAGCAT 7651
Db 650 AAGTTGCTGAAGAAATTAATGAAAAAGTAGATGAAGAAAGTAGCTGAAGAAATTAATGAAA 709
Qy 7652 ATCAAGTCTG---TAGAGCATTTGGAAATGAGAGCTGCTTATGCAAGATTAATTTCTA 7708
Db 710 AAGTGTCTGAAGAAATTAATGAAAAAGTAGATGAAGAAAGTAGCTGAAGAAATTAATTTGAAA 769
Qy 7709 ATGGAAGATCAAAATATCAGTATTAATAAATCTTAAGCTATTTGAAAAAATTTGATGTTA 7768
Db 770 AAGTGTCTGAAGAAATTAAGTAAAAAGTAGCTGAAGAAATTTGTTGAAAAAGTAGATGAAG 829
Qy 7769 TTTTAAAAAGATTAATTCGAAATTTGAGAGCGGAAAGCAAGATTAACCTGAGAGCGTAG 7828
Db 830 AAGTAGCTGAAGAAATTAAGTAAAAAGTAGATGAAGAAAGTAGCTGAAGAAAGTAGATCAAA 889
Qy 7829 CTGCGGAGCCATTAATCTCAAAAGCAAGAAATGAATGAATTCAGAGGTTGAAATTTGAA 7888
Db 890 AAGTAGAGAAAGAAAGTAACTGAAGAAATTAATGAAAAAGTAGATGAAGAAAGTAACTGTAAG 949
Qy 7889 AAGATATTTTGAATGAAGAAATTAAGATTAACCTAGCCCTTCTAAAGAAATTTGGAAGAGAA 7948
Db 950 AATTAAATGAAAAAGTAGATGAAGAAAGTGTGGAAGAAATTAATGAAAAAGTAGATGAAG 1009
Qy 7949 TCAATGTCAAAGTGGAAAAAGAAAAAGAGAGTACTGCTGAATCTCAAGAGAGCTTCTGTAG 8008
Db 1010 AAGTTGCTGAAGAAATTAATGAAAAAGTAGATGAATTAAGTAAAAAGTAGCTGAAG 1069
Qy 8009 GAGCAGTAGCAGGCGCAGGAATTAATTTCCGAAGCAAAAGATCCCGAAGCTTTATTGGA 8068
Db 1070 AATTAGTGAAGAAAGTAGATGAACAAGTAGCTGAAGAAATTAAGTAAAAAGTAGATGAAC 1129
Qy 8069 AAGTAGTCAAAAATCCGGAAGAAAGTATTTTTCATGCAATATATGTGAATATGAAAGCA 8128
Db 1130 AAGTAGCTGAAGAAATTAAGTAAAAAGTAGATGAACAAGTAGTGAAGAGTAGCTGAAG 1189
Qy 8129 CACATTAATGAAAGTAAACAGCAGTTTCTAAAGCAGTAAACAGTCTGTATTTGGAGAG 8188
Db 1190 AAGTAGCTGAAGAAAGTAGATGAAGAAAGTAGTGAAGAAAGTAGCTGAAGAGTAGCTGAAGAG 1249
Qy 8189 TTGAGATCACCAAGCAGAGAGCTACTGCTGCAAGTAAAACTATGTGTAAGTTGAGAGAG 8248
Db 1250 TAGCTGAAGAAAGTAGCTGAAGAAAGTAGCTGAAGAAAGTAGCTGAAGAAATTAAGTAAAAAG 1309

```

QY 8249 GAAATTTGTCAGAC---AAATCGATTGAATGCAATTTCTAAGTAGAAGGTTTGATG 8305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1310 TAGATGAAGAAAGTAGCTGAAAAAAGTAGTTGAAGAAAGAGTAGAAAAAGTAGAAGAG 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8306 AAGATAAAGTAAGTCTTAATCTTCTGTAGTATCAGAAATGAGAGGAAATTGCCGAG 8365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1370 TAGTTGAAGAAAGTAGATGAAGAGTAGCTGAAAAAAGTAGTTGAAGAAAGAGTAGAAAAAG 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8366 CAGAGTGAAATCTTCTACAGACAAAGTAATCTGAATCCGTAGTTGTTACGAAAGC 8425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1430 TACTTGAAGAAAGTAATTTGAAGAGTAGTTGAAGAAAGTAGCCGAAAGAGTAGCTGAAAAAG 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8426 AAGATTATGAAAAATATGATTACACAAAAAATATATTTCAGAG 8470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1490 TAGTTGAAGAAAGTAGAGTGAAAAAAGTAACAAAAATGATTTAATG 1534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: November 25, 2005, 14:17:05
 Job time : 31065 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2005, 22:13:25 ; Search time 3407 Seconds
(without alignments)

19025.771 Million cell updates/sec

Title: US-10-647-057-8

Perfect score: 9726

Sequence: 1 atgagcggcaccatacaaa.....gagaaaaaggaaaaaataga 9726

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_21.*
1: genebegn1980s:*
2: genebegn1990s:*
3: genebegn2000s:*
4: genebegn2001as:*
5: genebegn2001bs:*
6: genebegn2002as:*
7: genebegn2002bs:*
8: genebegn2003as:*
9: genebegn2003bs:*
10: genebegn2003cs:*
11: genebegn2003ds:*
12: genebegn2004as:*
13: genebegn2004bs:*
14: genebegn2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9726	100.0	9726	6	AA167640
2	9718	99.9	11130	6	AA167647
3	2780	28.6	2780	6	AA167642
4	2322	23.9	2322	6	AA167645
5	2141	22.0	2141	6	AA167643
6	1887	19.4	1887	6	AA167644
7	1130	11.6	1130	6	AA167641
8	116.2	1.2	5361	6	AA178868
9	116.2	1.2	5529	6	ABK50883
10	116.2	1.2	6152	2	AA178867
11	104.2	1.1	3399	2	AA105868
12	98.6	1.0	1701	14	AD272262
13	87.4	0.9	3579	3	AA170099
14	81.8	0.8	1686	2	AAQ87587
15	80.4	0.8	1300	12	ADP85917
16	79.2	0.8	4965	8	ADA89806
17	79.2	0.8	7434	4	AA552179
18	79.2	0.8	7434	8	AB114965
19	79.2	0.8	7434	8	ACF73459

20	79.2	0.8	7437	4	AA55232
21	77	0.8	204803	12	ADQ97348
22	76.4	0.8	4677	3	AA170259
23	76	0.8	1039	14	ACL64817
24	74.8	0.8	7104	4	AA551998
25	74.8	0.8	7107	4	AA554654
26	73.6	0.8	8155	2	AA174374
27	72.4	0.7	7446	8	ACA19708
28	72.4	0.7	7446	10	ACF05848
29	72.4	0.7	7446	13	ADU74046
30	68.8	0.7	1998	3	AA170212
31	68.8	0.7	31124	11	ACN44156
32	68.2	0.7	4911	8	AB115076
33	68.2	0.7	4997	8	ACA47636
34	67.6	0.7	4590	1	AA160472
35	67.6	0.7	7442	4	AA546686
36	66.2	0.7	15507	8	AA153680
37	65.8	0.7	1000	12	ADQ62833
38	65.8	0.7	1000	12	ADQ62832
39	65.4	0.7	5940	3	AA170105
40	65.2	0.7	2289	14	AD272266
41	64.4	0.7	1712	6	ABK50884
42	63.4	0.7	612	14	ACL64894
43	63.2	0.6	2407	8	ADA89805
44	63.2	0.6	3095	2	AAQ03875
45	63.2	0.6	6644	2	AA133181

ALIGNMENTS

RESULT 1	
AA167640	AA167640 standard; DNA; 9726 BP.
ID	
XX	AA167640;
AC	
XX	
DT	27-FEB-2002 (first entry)
XX	
DE	F. necrophorum leukotoxin protein encoding DNA.
XX	
KW	Leukotoxin; infection; immunisation; liver abscess; foot rot;
KW	bactericide; vaccine; ds.
XX	
OS	Fusobacterium necrophorum.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..9726
FT	/tag= a
FT	/product= "Leukotoxin"
XX	
PN	WO200180886-A2.
XX	
PD	01-NOV-2001.
XX	
XX	25-APR-2001; 2001WO-US013240.
XX	
PR	25-APR-2000; 2000US-00558257.
PR	24-APR-2001; 2001US-00841786.
PA	(UNITV) UNIV KANSAS STATE RES FOUND.
XX	
XX	Nagaraja TG, Stewart GC, Narayanan SK, Chengappa MM,
XX	WPI; 2002-049245/06.
XX	P-FSDB; AA666005.
XX	
PT	Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an
PT	animal against an infection e.g. foot rot, or liver abscesses caused by
PT	the bacterium.
XX	
PS	Claim 10; Page 90-95; 108pp; English.
XX	

CC The invention provides an isolated *Fusobacterium necrophorum* leukotoxin
CC polypeptide and its truncated versions. The leukotoxin protein can be
CC expressed by standard recombinant methodology. The leukotoxin gene is
CC useful for preparing a vaccine which confers effective immunity against
CC infection caused by *F. necrophorum*. The vaccine is useful for immunising
CC an animal against liver abscesses caused by *F. necrophorum* and for
CC preventing foot rot caused by *F. necrophorum* infection. The present
CC sequence represents a DNA encoding a *F. necrophorum* full-length
CC leukotoxin polypeptide
XX

Sequence 9726 BP; 3678 A; 1304 C; 2238 G; 2506 T; 0 U; 0 Other;

Query Match 100.0%; Score 9726; DB 6; Length 9726;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGACGGGATCAAAATTAAGTTTCAGAGACAGAGAGAGATTCAGATTCTAAAAA 60
DB 1 ATGACGGGATCAAAATTAAGTTTCAGAGACAGAGAGAGATTCAGATTCTAAAAA 60
QY 61 GTTTTAATGATTTGGGATTTGATTAACATATGAGGCTTAAGGCTTAATCAATC 120
DB 61 GTTTTAATGATTTGGGATTTGATTAACATATGAGGCTTAATCAATC 120
QY 121 ACCGGGACTGAGAAATTTTGGAA CAAAAATAGAAAAAAGATATGTTTATGACATTACT 180
DB 121 ACCGGGACTGAGAAATTTTGGAA CAAAAATAGAAAAAAGATATGTTTATGACATTACT 180
QY 181 ACAAACAAGATTCAGAGGAGAACGCTTTAACAGTTTAATAGATTGGCTTTAACAGAA 240
DB 181 ACAAACAAGATTCAGAGGAGAACGCTTTAACAGTTTAATAGATTGGCTTTAACAGAA 240
QY 241 AATAATATAGCAATCTATATTTGGGAAAAAGAAATAGACGGGGTAAATATCTTTT 300
DB 241 AATAATATAGCAATCTATATTTGGGAAAAAGAAATAGACGGGGTAAATATCTTTT 300
QY 301 AACTTTGTCAATGAGAAAAATTAAGTAGATGGATTATCAAGAAATTCAGAAAAATAA 360
DB 301 AACTTTGTCAATGAGAAAAATTAAGTAGATGGATTATCAAGAAATTCAGAAAAATAA 360
QY 361 ATTGAGAGAAATTTATTTCTTAAGCTCGAAAGGATGGCATGAGAAAAATGAGATT 420
DB 361 ATTGAGAGAAATTTATTTCTTAAGCTCGAAAGGATGGCATGAGAAAAATGAGATT 420
QY 421 ATCAATGCTGGTCTTTTCATTCTATATTCGAAAAACAAGATTTTAAAGAGCTTGG 480
DB 421 ATCAATGCTGGTCTTTTCATTCTATATTCGAAAAACAAGATTTTAAAGAGCTTGG 480
QY 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGAAATCATTCAGATGATGAAAGATAAA 540
DB 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGAAATCATTCAGATGATGAAAGATAAA 540
QY 541 ATTCCATTGAATCCGAATGGAAGCATTCAGGTAGAAAGAAAAATCAATGCTGTTAAAGC 600
DB 541 ATTCCATTGAATCCGAATGGAAGCATTCAGGTAGAAAGAAAAATCAATGCTGTTAAAGC 600
QY 601 ATCGGTTTATATCGCGCGGATTTAGATTGAAGAATCTGCAATCTAAAGCAGGAATT 660
DB 601 ATCGGTTTATATCGCGCGGATTTAGATTGAAGAATCTGCAATCTAAAGCAGGAATT 660
QY 661 ACAGATTTTAAAAATTTAGTCAATATTAAGTAGATCGAATTAATTCTGGCTGACGGAGAT 720
DB 661 ACAGATTTTAAAAATTTAGTCAATATTAAGTAGATCGAATTAATTCTGGCTGACGGAGAT 720
QY 721 TTTAAAGCTACCAAGACAAATCTGAGATATTTATTTCTTCACTCAATAGATTCTCCT 780
DB 721 TTTAAAGCTACCAAGACAAATCTGAGATATTTATTTCTTCACTCAATAGATTCTCCT 780
QY 781 CAAAAAGCTATGGAAAAAATTCACCTGTGGAAGAGAAATGAAAGAAATGTTAAAAAGAA 840
DB 781 CAAAAAGCTATGGAAAAAATTCACCTGTGGAAGAGAAATGAAAGAAATGTTAAAAAGAA 840
QY 841 AATACCAAGCAAAATATGAAATCTGATGCTGTATTGGAAGCAGATGGAATTAATAAATT 900
```

```
DB 841 AATACCAAGCAAAATATGAAATCTGATCTGTATTGGAAGCAGATGGAATTAATAAATT 900
QY 901 AGTGGAAAGCTACCAATGGAGATTATTAAGAAAGAGGGAAAAAGAACTTATTAAC 960
DB 901 AGTGGAAAGCTACCAATGGAGATTATTAAGAAAGAGGGAAAAAGAACTTATTAAC 960
QY 961 ACTCCTTAATGTTTATCAGATGTGAGAGCTCCGTAAGATTAATTAAGAAAGATCTATA 1020
DB 961 ACTCCTTAATGTTTATCAGATGTGAGAGCTCCGTAAGATTAATTAAGAAAGATCTATA 1020
QY 1021 GAAAAAGATGTGACATTACAGCTGAGCAAAAGAAATTTCTATGATGCAACTTAACTACT 1080
DB 1021 GAAAAAGATGTGACATTACAGCTGAGCAAAAGAAATTTCTATGATGCAACTTAACTACT 1080
QY 1081 AAGCTTGCAAAAGCACTCTTTTAGCTTTGTTCAGAGTTCTATTTCTCTATCAATTTAAT 1140
DB 1081 AAGCTTGCAAAAGCACTCTTTTAGCTTTGTTCAGAGTTCTATTTCTCTATCAATTTAAT 1140
QY 1141 GGAATTTTAGGTTTATGACAAGTAAGTCAAGTGCATTAATTGAAAAAGATGCCAAAGTC 1200
DB 1141 GGAATTTTAGGTTTATGACAAGTAAGTCAAGTGCATTAATTGAAAAAGATGCCAAAGTC 1200
QY 1201 GAAGCAACAGAAAGAAAGCAAAATATTCATTCTTACAGTGAATAGACAACTATGGGA 1260
DB 1201 GAAGCAACAGAAAGAAAGCAAAATATTCATTCTTACAGTGAATAGACAACTATGGGA 1260
QY 1261 GCAGCTACTTCTCCCTTAAAAATTCACATTTATTTTGGAGAAAGCCAAATGAAAACCTT 1320
DB 1261 GCAGCTACTTCTCCCTTAAAAATTCACATTTATTTTGGAGAAAGCCAAATGAAAACCTT 1320
QY 1321 CTCAATATCGAGAGCGGATATATTTCTGCAAAAGTAATTCGAATGTAATTAAGAAAGA 1380
DB 1321 CTCAATATCGAGAGCGGATATATTTCTGCAAAAGTAATTCGAATGTAATTAAGAAAGA 1380
QY 1381 GAAGTAAATCGAAGGAGAAAGCAGATATTAATCTTCAAAATCTGAAAATCTATATGATGCT 1440
DB 1381 GAAGTAAATCGAAGGAGAAAGCAGATATTAATCTTCAAAATCTGAAAATCTATATGATGCT 1440
QY 1441 TCTGTTTCTGTTGGAGCATGAGAGATTCGAATTAAGTCTCTTCAGTATGGTGAAG 1500
DB 1441 TCTGTTTCTGTTGGAGCATGAGAGATTCGAATTAAGTCTCTTCAGTATGGTGAAG 1500
QY 1501 GAAGGAGAAATTAATCTTCGCTCAAGATTCGTAAGAGGCAAAAGTGAATCAGAAAGC 1560
DB 1501 GAAGGAGAAATTAATCTTCGCTCAAGATTCGTAAGAGGCAAAAGTGAATCAGAAAGC 1560
QY 1561 GATGATGTAATGTGAGAAAGTGAAGCGATTAATTCATTCGAGCTGCTGTAAAGGTGGA 1620
DB 1561 GATGATGTAATGTGAGAAAGTGAAGCGATTAATTCATTCGAGCTGCTGTAAAGGTGGA 1620
QY 1621 TTGGGGGATATGATGTAATGGGGTGTGGCTGCAAAATATTTCTAATTAATGCTTCCTC 1680
DB 1621 TTGGGGGATATGATGTAATGGGGTGTGGCTGCAAAATATTTCTAATTAATGCTTCCTC 1680
QY 1681 CGTATAGATGATGATGATATCTACATGCAAGCAAGAGCACTAAATGTGAGAGCTCATAC 1740
DB 1681 CGTATAGATGATGATGATATCTACATGCAAGCAAGAGCACTAAATGTGAGAGCTCATAC 1740
QY 1741 ATTAATAAAATAGTGTTCGCAAAACAGATCTGATTTGGAACTTCCAACTTATAGAT 1800
DB 1741 ATTAATAAAATAGTGTTCGCAAAACAGATCTGATTTGGAACTTCCAACTTATAGAT 1800
QY 1801 GATCAGCTTTATGAATCAGGTGATCTAAATCAATTTTATGATGCAATAAACAGCGGTTT 1860
DB 1801 GATCAGCTTTATGAATCAGGTGATCTAAATCAATTTTATGATGCAATAAACAGCGGTTT 1860
QY 1861 GGAAGAGCAGTGTCAATGAGAAATTAAGAAATTAAGCAAGAACTTATAGTGGCT 1920
DB 1861 GGAAGAGCAGTGTCAATGAGAAATTAAGAAATTAAGCAAGAACTTATAGTGGCT 1920
QY 1921 GTGTCTGCAACCAATAGCAAAATCATTAATATCTGCTTCTGTGGCAATGAGAGAGTGA 1980
```

Db 1921 GTGCTGCAACCATAGCAATCATTAATTCGCTTCTGTGGCAATAGAGAGAGTGA 1980
Qy 1981 AGACTTTCTTCAAGAGTGGAAAGGAGTAATGTAAGGGCAATTAATGAAGCTCAAAATCTT 2040
Db 1981 AGACTTTCTTCAAGAGTGGAAAGGAGTAATGTAAGGGCAATTAATGAAGCTCAAAATCTT 2040
Qy 2041 CGAGGACCTAGCTCAAGTGAAGTGGCTGTACAAAGAAAGAAAGAAAGAAAGAAAGCTTAAT 2100
Db 2041 CGAGGACCTAGCTCAAGTGAAGTGGCTGTACAAAGAAAGAAAGAAAGAAAGAAAGCTTAAT 2100
Qy 2101 GGAATGCAAGCAGTTTTTATGAAACTATATAAAATTAATGCTTCTGTGCAATTCGCAAT 2160
Db 2101 GGAATGCAAGCAGTTTTTATGAAACTATATAAAATTAATGCTTCTGTGCAATTCGCAAT 2160
Qy 2161 CATGCTGAATGGTATCGGAAGGAAATTTGATATCAACAGTGAATTAATTAATTAAT 2220
Db 2161 CATGCTGAATGGTATCGGAAGGAAATTTGATATCAACAGTGAATTAATTAATTAAT 2220
Qy 2221 AAAAATCCTTCAAAAATGGCAAAAGTCTGTATTAATTAATTAATTAATTAATTAAT 2280
Db 2221 AAAAATCCTTCAAAAATGGCAAAAGTCTGTATTAATTAATTAATTAATTAATTAAT 2280
Qy 2281 TTGGAAGAAAGAAACGAAATCTCAGATATGATCCGAAAGATTTGAATCTATTGAAAA 2340
Db 2281 TTGGAAGAAAGAAACGAAATCTCAGATATGATCCGAAAGATTTGAATCTATTGAAAA 2340
Qy 2341 TTATGAAATGCAATTTTCAGAAAAATTTGATGAAAAACCGAGCTTTTACTAAATGGTGA 2400
Db 2341 TTATGAAATGCAATTTTCAGAAAAATTTGATGAAAAACCGAGCTTTTACTAAATGGTGA 2400
Qy 2401 AGAATGCAATTAATCTTCCGGAATGGAATTTGAAAAACGGAACGCTATTAATTAATGCA 2460
Db 2401 AGAATGCAATTAATCTTCCGGAATGGAATTTGAAAAACGGAACGCTATTAATTAATGCA 2460
Qy 2461 AACTATGTTCAAGGAGAAATGAAAAATTTAGAGAAAAATTTACCGAAAGATTTAAAGCT 2520
Db 2461 AACTATGTTCAAGGAGAAATGAAAAATTTAGAGAAAAATTTACCGAAAGATTTAAAGCT 2520
Qy 2521 TTTTCAGAGAGATGAGTGAATTAAGAAATTTGAATTTTACAGAGATGAGAAAT 2580
Db 2521 TTTTCAGAGAGATGAGTGAATTAAGAAATTTGAATTTTACAGAGATGAGAAAT 2580
Qy 2581 TATGCAATTTTCAACCTTTTACCTTCCGAGCTTAATGAGAAAGAAATGCTTCTTCT 2640
Db 2581 TATGCAATTTTCAACCTTTTACCTTCCGAGCTTAATGAGAAAGAAATGCTTCTTCT 2640
Qy 2641 GTGGAGAGAGCTGTTTCTGTGGTGAACAGAGAAATTTATGCAAGGATATCCGTTGAAAA 2700
Db 2641 GTGGAGAGAGCTGTTTCTGTGGTGAACAGAGAAATTTATGCAAGGATATCCGTTGAAAA 2700
Qy 2701 GGAGCTAAATCTGTGCAAAAAAAGATTTAAATTAATTAATTAATTAATTAATTAATTA 2760
Db 2701 GGAGCTAAATCTGTGCAAAAAAAGATTTAAATTAATTAATTAATTAATTAATTAATTA 2760
Qy 2761 GTGAATTTATGTTGAAAAATTTGAGCTTGGAGAGCAATCAATCCGAAAGTGCAGCTGA 2820
Db 2761 GTGAATTTATGTTGAAAAATTTGAGCTTGGAGAGCAATCAATCCGAAAGTGCAGCTGA 2820
Qy 2821 GGAAGATTAAATGTTGAAAAATTTGAGCTATGCTATGAGAACTAAAGAAAGAAAGCT 2880
Db 2821 GGAAGATTAAATGTTGAAAAATTTGAGCTATGCTATGAGAACTAAAGAAAGAAAGCT 2880
Qy 2881 GAATTTATCAGAGAAAAATTAATGCAATGCAATTTGAAAGAAAGCTTTTCAATGAGGGA 2940
Db 2881 GAATTTATCAGAGAAAAATTAATGCAATGCAATTTGAAAGAAAGCTTTTCAATGAGGGA 2940
Qy 2941 TCTTTTAAATGTTGCTCAGGTGGAAATTCGAATCAATGGAATGGAAAGTTATATGAGAGGT 3000
Db 2941 TCTTTTAAATGTTGCTCAGGTGGAAATTCGAATCAATGGAATGGAAAGTTATATGAGAGGT 3000
Qy 3001 ATCAGTAAGGCAAGATTTTCAATGATGACGAAAGCATATTTGAAGCTAATTAATTAATTA 3060
Db 3001 ATCAGTAAGGCAAGATTTTCAATGATGACGAAAGCATATTTGAAGCTAATTAATTAATTA 3060

Qy 3061 GCTTTAAACAGTAAGAAATGATACTTCTGTGTAATGCAATGCTGCCGGTTACGCCGAATCCGA 3120
Db 3061 GCTTTAAACAGTAAGAAATGATACTTCTGTGTAATGCAATGCTGCCGGTTACGCCGAATCCGA 3120
Qy 3121 ACGAAAAATGCGCGGTGCGGGGTGCTGTTCGGTAAATGATATGATATTTCAACAAA 3180
Db 3121 ACGAAAAATGCGCGGTGCGGGGTGCTGTTCGGTAAATGATATGATATTTCAACAAA 3180
Qy 3181 GCTTCCATTTGAAGTAATGACGAAGAGCAAAATTAATGATTAAGAAATTAAGATGATGA 3240
Db 3181 GCTTCCATTTGAAGTAATGACGAAGAGCAAAATTAATGATTAAGAAATTAAGATGATGA 3240
Qy 3241 GTAACAGTACTGCGGAATCTTTAGAAATGATGCAAAAAAGACCGGAACAATCAACAGT 3300
Db 3241 GTAACAGTACTGCGGAATCTTTAGAAATGATGCAAAAAAGACCGGAACAATCAACAGT 3300
Qy 3301 ATTTCTGTTGCGGAGAAATTAATAGGTGGAAGTAAACCGAGTGAAGAAAAACCGAAA 3360
Db 3301 ATTTCTGTTGCGGAGAAATTAATAGGTGGAAGTAAACCGAGTGAAGAAAAACCGAAA 3360
Qy 3361 TCAGAAAGAAAGCCAGAGGATTTTGGCAAAATTCGAAAAACAAGTGAATCTGTGAAAA 3420
Db 3361 TCAGAAAGAAAGCCAGAGGATTTTGGCAAAATTCGAAAAACAAGTGAATCTGTGAAAA 3420
Qy 3421 AATTAATTTAGGATATGATGATTCATTAACAGAAAAATTAACAAATTAATTAATTAAT 3480
Db 3421 AATTAATTTAGGATATGATGATTCATTAACAGAAAAATTAACAAATTAATTAATTAAT 3480
Qy 3481 GGAGTAAATTAAGAAACCGGGAATCTTCTCGAATGCTTTCTCATATCTCCGATTAAGAGCG 3540
Db 3481 GGAGTAAATTAAGAAACCGGGAATCTTCTCGAATGCTTTCTCATATCTCCGATTAAGAGCG 3540
Qy 3541 TCTTCAAGTTGGAGCTTCTGGAAGTCTTTTCAATTAATTAATTAATTAATTAATTAAT 3600
Db 3541 TCTTCAAGTTGGAGCTTCTGGAAGTCTTTTCAATTAATTAATTAATTAATTAATTAAT 3600
Qy 3601 GCTGTCGTAAGTGAAGTAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3660
Db 3601 GCTGTCGTAAGTGAAGTAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3660
Qy 3661 TCTTCTGATTTCACTTTTGTGAGCAATGAGGCGGATCTGCTGCACTTCAATGATCAT 3720
Db 3661 TCTTCTGATTTCACTTTTGTGAGCAATGAGGCGGATCTGCTGCACTTCAATGATCAT 3720
Qy 3721 ATTTGAAAGTGAAGTAAGCAATCACTGCTGCTTTAGCTGGAAGGCTGCTGAATTAAT 3780
Db 3721 ATTTGAAAGTGAAGTAAGCAATCACTGCTGCTTTAGCTGGAAGGCTGCTGAATTAAT 3780
Qy 3781 ATTTCAAGTAAACCAAGTCTTTGTTGAAATTAATTAATTAATTAATTAATTAATTA 3840
Db 3781 ATTTCAAGTAAACCAAGTCTTTGTTGAAATTAATTAATTAATTAATTAATTAATTA 3840
Qy 3841 AAAAGTAAATGCTTTGAGTGAAGAACTCAAGTGAAGCAAGGAGCAAGTGTGAAGCAAT 3900
Db 3841 AAAAGTAAATGCTTTGAGTGAAGAACTCAAGTGAAGCAAGGAGCAAGTGTGAAGCAAT 3900
Qy 3901 AAAAGTGAAGGAGCAAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3960
Db 3901 AAAAGTGAAGGAGCAAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3960
Qy 3961 GTGAACAAATGAAGTTTCTGCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
Db 3961 GTGAACAAATGAAGTTTCTGCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
Qy 4021 CAAAAAATGATGTTGATGCTGCTTATCAAGCGGACCAACCAATGACAGAGCTTTTA 4080
Db 4021 CAAAAAATGATGTTGATGCTGCTTATCAAGCGGACCAACCAATGACAGAGCTTTTA 4080
Qy 4081 AATTTTCAAGCTGGAAGATCAATGGAATGTAAGGGCTATGCTGATCTGTTGCCAAATTA 4140
Db 4081 AATTTTCAAGCTGGAAGATCAATGGAATGTAAGGGCTATGCTGATCTGTTGCCAAATTA 4140

QY 4141 AACAAAGTAATGCTTCTATTAGTGTGGAGATATACTAACGTTAATGACGGAC 4200
Db 4141 AACACAAAGTAATGCTTCTATTAGTGTGGAGATATACTAACGTTAATGACGGAC 4200
QY 4201 GCAAAAGCTCTTTTAGCAACCACTCAAGTGAATGCTGCAAGTGAACGAGAGGACAAAT 4260
Db 4201 GCAAAAGCTCTTTTAGCAACCACTCAAGTGAATGCTGCAAGTGAACGAGAGGACAAAT 4260
QY 4261 AGTTCTGGAGCGGAGTTAGAAATATCAAGGGGCTGTTCTGCAATTAAGTTGACAT 4320
Db 4261 AGTTCTGGAGCGGAGTTAGAAATATCAAGGGGCTGTTCTGCAATTAAGTTGACAT 4320
QY 4321 GACGTGAAGCTAGCGTGTATTAATCTTCATCGAAGAGCTAATGAATCAATGTCAT 4380
Db 4321 GACGTGAAGCTAGCGTGTATTAATCTTCATCGAAGAGCTAATGAATCAATGTCAT 4380
QY 4381 GCCAAAGTGTCAAGGAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAATGGA 4440
Db 4381 GCCAAAGTGTCAAGGAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAATGGA 4440
QY 4441 AAAAGTAAAAAATATTTAGAAAGATGTGTGATTAATAGACTGGAATGTTAATACG 4500
Db 4441 AAAAGTAAAAAATATTTAGAAAGATGTGTGATTAATAGACTGGAATGTTAATACG 4500
QY 4501 AAGGAACAATAAGAAAAAGCAAGAAAAAGAGAGCGTCAATTTGTAATGCTGCTTTA 4560
Db 4501 AAGGAACAATAAGAAAAAGCAAGAAAAAGAGAGCGTCAATTTGTAATGCTGCTTTA 4560
QY 4561 TCGGTGCTGGAAAGGATTAATCCGCTGAGAGAGTACGTAATGCAATCTGTTAA 4620
Db 4561 TCGGTGCTGGAAAGGATTAATCCGCTGAGAGAGTACGTAATGCAATCTGTTAA 4620
QY 4621 AATAAATTTAAAGCAAGATTGAGTGAAGCAATAAGAAAGCGAGAGATTAATATAT 4680
Db 4621 AATAAATTTAAAGCAAGATTGAGTGAAGCAATAAGAAAGCGAGAGATTAATATAT 4680
QY 4681 GCGAAACATGTAAATGTGAGAGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGA 4740
Db 4681 GCGAAACATGTAAATGTGAGAGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGA 4740
QY 4741 CTGCTATACAGAAAGATGCTTTTTCAGGAATGGGATCTGAGCATGCGACAGACTTAACA 4800
Db 4741 CTGCTATACAGAAAGATGCTTTTTCAGGAATGGGATCTGAGCATGCGACAGACTTAACA 4800
QY 4801 AATGACAGATTTGCAAGAGTGAATAAGGAAGAAATTTCTGCTGATTTCTTAATGTAAC 4860
Db 4801 AATGACAGATTTGCAAGAGTGAATAAGGAAGAAATTTCTGCTGATTTCTTAATGTAAC 4860
QY 4861 GCAAAATTAATTCATTTCTTGGGCTGAATGTTGGGGAACCAATTCGCGGTTCTTTCTACG 4920
Db 4861 GCAAAATTAATTCATTTCTTGGGCTGAATGTTGGGGAACCAATTCGCGGTTCTTTCTACG 4920
QY 4921 GCGGTAGAGCTGCTTTTTCGCAATATCTCTTCAATAAACCCTGCTTGTGATACA 4980
Db 4921 GCGGTAGAGCTGCTTTTTCGCAATATCTCTTCAATAAACCCTGCTTGTGATACA 4980
QY 4981 GAAAGCAAGGTAAATCCTTTAGTGAAGAAAGATTAAGAAAGCAATGTCAGGCTTTGAT 5040
Db 4981 GAAAGCAAGGTAAATCCTTTAGTGAAGAAAGATTAAGAAAGCAATGTCAGGCTTTGAT 5040
QY 5041 GATTCATATTAACAACGTTTCTGCTGAGGCGCTGCAAGTATTAACAGAGCTGGAATC 5100
Db 5041 GATTCATATTAACAACGTTTCTGCTGAGGCGCTGCAAGTATTAACAGAGCTGGAATC 5100
QY 5101 GAGAGATGTATCTGTCAATCTGTGTTCTGATGAACGAAAGCTTTAGTATGATCT 5160
Db 5101 GAGAGATGTATCTGTCAATCTGTGTTCTGATGAACGAAAGCTTTAGTATGATCT 5160
QY 5161 GAGTTGAAGAGTAAAGTTCTTTCAATGATGAAGCAAAAGCAATTAATTAATCA 5220
Db 5161 GAGTTGAAGAGTAAAGTTCTTTCAATGATGAAGCAAAAGCAATTAATTAATCA 5220
QY 5221 ATTGCCGAAATGCAAAATGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACA 5280

Db 5221 ATTGCCGAAATGCAAAATGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACA 5280
QY 5281 AATATGGAAGAAACATCTGTTATAGCTTATGTAAGAAACAGTAAATTAACAAGCGCAAT 5340
Db 5281 AATATGGAAGAAACATCTGTTATAGCTTATGTAAGAAACAGTAAATTAACAAGCGCAAT 5340
QY 5341 GATCAAGATAGAAAAATATCAATGTGACTCAAAAGATTTATATGACCAATACATATA 5400
Db 5341 GATCAAGATAGAAAAATATCAATGTGACTCAAAAGATTTATATGACCAATACATATA 5400
QY 5401 GCAGTCGAGTTGAGAGAGCAAAAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTAC 5460
Db 5401 GCAGTCGAGTTGAGAGAGCAAAAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTAC 5460
QY 5461 TTGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTTAGCAAGATTTAGAGGA 5520
Db 5461 TTGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTTAGCAAGATTTAGAGGA 5520
QY 5521 GAAATTAATGGAATTAAGAAAGGCAATGTTAATGTTCTAGAGTGAATTAACGAGTCA 5580
Db 5521 GAAATTAATGGAATTAAGAAAGGCAATGTTAATGTTCTAGAGTGAATTAACGAGTCA 5580
QY 5581 GTGTCACAAAATGCGACACAGTCTTCCGAGCAAGTGAACAAGCTGCAAGTGAAGCTGGA 5640
Db 5581 GTGTCACAAAATGCGACACAGTCTTCCGAGCAAGTGAACAAGCTGCAAGTGAAGCTGGA 5640
QY 5641 GTAGCATTAATTAATTAATCACAAAATCTTCTGCACTATTAATAAATAGTCTAAAT 5700
Db 5641 GTAGCATTAATTAATTAATCACAAAATCTTCTGCACTATTAATAAATAGTCTAAAT 5700
QY 5701 GTACAGATTAATTAATTAATCACAAAATCTTCTGCACTATTAATAAATAGTCTAAAT 5760
Db 5701 GTACAGATTAATTAATTAATCACAAAATCTTCTGCACTATTAATAAATAGTCTAAAT 5760
QY 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGTGAACAGTTCTGTAGCAGTGAATTAAGTTGTA 5820
Db 5761 GCTGAGTTGAGAGAGCTGAGAGAGCTGAGTGAACAGTTCTGTAGCAGTGAATTAAGTTGTA 5820
QY 5821 AATAATACATAGACGAATTAATTAATCAAGCAAAATCACTGCGAAGGAAATGTGGAAT 5880
Db 5821 AATAATACATAGACGAATTAATTAATCAAGCAAAATCACTGCGAAGGAAATGTGGAAT 5880
QY 5881 ATTAAGAGCTGATGAGCGGTAAATGCTAATATGAGGAACAGTGTGAGAGTGGCCGT 5940
Db 5881 ATTAAGAGCTGATGAGCGGTAAATGCTAATATGAGGAACAGTGTGAGAGTGGCCGT 5940
QY 5941 GCAGCAATAGAGGCTCAACCAAGTGTGAATGAATTAACAGATCTCAAAAGCAATATGTA 6000
Db 5941 GCAGCAATAGAGGCTCAACCAAGTGTGAATGAATTAACAGATCTCAAAAGCAATATGTA 6000
QY 6001 AAAAGTTTACAGATGTTGCTTAAGAGAAACAGATGTTATTAATCTACCAAGGCGCA 6060
Db 6001 AAAAGTTTACAGATGTTGCTTAAGAGAAACAGATGTTATTAATCTACCAAGGCGCA 6060
QY 6061 GTAGATAAGGTAGATTAAGTATTCAAAATCTTAATTAATTAAGAGACTTAATCAAA 6120
Db 6061 GTAGATAAGGTAGATTAAGTATTCAAAATCTTAATTAATTAAGAGACTTAATCAAA 6120
QY 6121 AAAAGAAAAATAGTAATTAAGAAAAAGATTTGTTACCAATAGTCACTACTCATCTTTA 6180
Db 6121 AAAAGAAAAATAGTAATTAAGAAAAAGATTTGTTACCAATAGTCACTACTCATCTTTA 6180
QY 6181 AAATCTTATTTAGCAAAATGCGGCTGTTCAAGACAGCGGAGTGGCAAACTGTAT 6240
Db 6181 AAATCTTATTTAGCAAAATGCGGCTGTTCAAGACAGCGGAGTGGCAAACTGTAT 6240
QY 6241 ATCAACAAGGTTTATGAGAAACAGAGCTGTGTAAGAAATCTATTAATTAAGCAAAA 6300
Db 6241 ATCAACAAGGTTTATGAGAAACAGAGCTGTGTAAGAAATCTATTAATTAAGCAAAA 6300
QY 6301 CATTAATCTGTAATATCAGAGATTAACAGATTAATCAATCGAGTGAATGAGTTCTGTGAT 6360

Dh 6301 CATTATCTGTAATAATCAGAGATTACAGAAATTCATTCGAGTAGTAGTTCGTGTGTT 6360
Qy 6361 GTTGTGGAATATGAGAGTAGAGCTTCTTGATACCAATATTTATTAAGAAATACC 6420
Db 6361 GTTGTGGAATATGAGAGTAGAGCTTCTTGATACCAATATTTATTAAGAAATACC 6420
Qy 6421 AAGAAGAGAGTTGAAAAAATCAATGCTGATGAAAGTTTCGAGAGAAAGCTGAAATTT 6480
Db 6421 AAGAAGAGAGTTGAAAAAATCAATGCTGATGAAAGTTTCGAGAGAAAGCTGAAATTT 6480
Qy 6481 ACAGCAGATTCTAAGCAAGAAATTTCTTTTGAGAGTCGAGAGCCGAGGTA 6540
Db 6481 ACAGCAGATTCTAAGCAAGAAATTTCTTTTGAGAGTCGAGAGCCGAGGTA 6540
Qy 6541 GGAGCCGAGTAGCGAGAAACCGTTTCCGTAATTCATTTGACAGAAAGCGAGAGTAGAT 6600
Db 6541 GGAGCCGAGTAGCGAGAAACCGTTTCCGTAATTCATTTGACAGAAAGCGAGAGTAGAT 6600
Qy 6601 GTGGAAGAAAGCAAAATTTTGTGTAAGAAAGCTGAGATTACAGCAAAACGTTATAGTTCT 6660
Db 6601 GTGGAAGAAAGCAAAATTTTGTGTAAGAAAGCTGAGATTACAGCAAAACGTTATAGTTCT 6660
Qy 6661 GTTGCAATTTGAAATGCGCAGTCGAGTCGCTGCAAAAAGAGCTGAAATTTGAGCAGCA 6720
Db 6661 GTTGCAATTTGAAATGCGCAGTCGAGTCGCTGCAAAAAGAGCTGAAATTTGAGCAGCA 6720
Qy 6721 GTGCAAGTTTACCAAGATGATCAAAACACGAGAGCAAGTAGAATAATTTCTAATAATATG 6780
Db 6721 GTGCAAGTTTACCAAGATGATCAAAACACGAGAGCAAGTAGAATAATTTCTAATAATATG 6780
Qy 6781 ACTGCAAAACAATTGATGATATAGCAGAAATGAGATAAATCGATATCGATATCGAT 6840
Db 6781 ACTGCAAAACAATTGATGATATAGCAGAAATGAGATAAATCGATATCGATATCGAT 6840
Qy 6841 TCAGCCGAGAGTCGAAATTTCTGACGCGGAGTATCTGAGAGTGTCTCTCATATATAT 6900
Db 6841 TCAGCCGAGAGTCGAAATTTCTGACGCGGAGTATCTGAGAGTGTCTCTCATATATAT 6900
Qy 6901 GCAAAATAGGTAGAAACAGATATCGATCATAGTACTTTACACTTTCTTCACTGATATAT 6960
Db 6901 GCAAAATAGGTAGAAACAGATATCGATCATAGTACTTTACACTTTCTTCACTGATATAT 6960
Qy 6961 GTTAAAGCTCTTAAATTAATTTTCTGATTCGATGACGCGGTGAGAGCCGCAAGTCTT 7020
Db 6961 GTTAAAGCTCTTAAATTAATTTTCTGATTCGATGACGCGGTGAGAGCCGCAAGTCTT 7020
Qy 7021 GCAGCAGTTACCGAGAGTGTCTCTGTTAACAATAATAGTTCTGTAGTACGAGTT 7080
Db 7021 GCAGCAGTTACCGAGAGTGTCTCTGTTAACAATAATAGTTCTGTAGTACGAGTT 7080
Qy 7081 CACAATTAACCTGATTTGACTTCCGTACGAGAAAGAAAGTAATGTAAACGCAAAAGAGAA 7140
Db 7081 CACAATTAACCTGATTTGACTTCCGTACGAGAAAGAAAGTAATGTAAACGCAAAAGAGAA 7140
Qy 7141 AAAAAATATTAAGCAACACAGCAAAATGCAAGAAATGCGAGAGCAATTCGAGGCAAT 7200
Db 7141 AAAAAATATTAAGCAACACAGCAAAATGCAAGAAATGCGAGAGCAATTCGAGGCAAT 7200
Qy 7201 GTCTTGATTAATTAATTTGAAACAGCTGTAGAGATAGAAATAATTTCTGAGAGAAAGGA 7260
Db 7201 GTCTTGATTAATTAATTTGAAACAGCTGTAGAGATAGAAATAATTTCTGAGAGAAAGGA 7260
Qy 7261 ACAGAGTTTAAAAAATTGAGAGCAAGTTAACAAAGAAACAGATAAAAAGTAATGAT 7320
Db 7261 ACAGAGTTTAAAAAATTGAGAGCAAGTTAACAAAGAAACAGATAAAAAGTAATGAT 7320
Qy 7321 GCTACGAAAAAAATCTTACATCAGCAGATATTTCTACAGAGATATCTTCTGTAAGAGG 7380
Db 7321 GCTACGAAAAAAATCTTACATCAGCAGATATTTCTACAGAGATATCTTCTGTAAGAGG 7380
Qy 7381 GATAGAGAGATATCTACGAGGAGAGAAATTAAGCAATTTGAGAGCTTCTGATATATAT 7440
Db 7381 GATAGAGAGATATCTACGAGGAGAGAAATTAAGCAATTTGAGAGCTTCTGATATATAT 7440

Qy 7441 GGAAGAAATGATGATATTTACAGAGAGCAAGAAATTAATATCATTTACTGAGTGTGTTG 7500
Db 7441 GGAAGAAATGATGATATTTACAGAGAGCAAGAAATTAATATCATTTACTGAGTGTGTTG 7500
Qy 7501 GGAATGCAAGGCTTGTCTCCGATCAGGAAACAGTGGCAGTTACAAATATTTAAAAAGAAAT 7560
Db 7501 GGAATGCAAGGCTTGTCTCCGATCAGGAAACAGTGGCAGTTACAAATATTTAAAAAGAAAT 7560
Qy 7561 TCCGAGTTACTGTTGAAAAATTTCTTTGTGAAAGAGAGCTGAAAAAGTAATGTTAGATCG 7620
Db 7561 TCCGAGTTACTGTTGAAAAATTTCTTTGTGAAAGAGAGCTGAAAAAGTAATGTTAGATCG 7620
Qy 7621 GATATTTACAGAAATGTTGCTTTTAAACAGCATATCAAGGCTCTGAGAGCATTTGGAATA 7680
Db 7621 GATATTTACAGAAATGTTGCTTTTAAACAGCATATCAAGGCTCTGAGAGCATTTGGAATA 7680
Qy 7681 GGAGCTGCTATGCAAGATTTAAATTTCTAATGAAAGATCAAAATATCAGATTTAAAAATCT 7740
Db 7681 GGAGCTGCTATGCAAGATTTAAATTTCTAATGAAAGATCAAAATATCAGATTTAAAAATCT 7740
Qy 7741 AAGCTATTAAGAAAAATATGATGTTATTTGTAAGATTAATCGGAATTTGAGACCGGAA 7800
Db 7741 AAGCTATTAAGAAAAATATGATGTTATTTGTAAGATTAATCGGAATTTGAGACCGGAA 7800
Qy 7801 GCAAAAGAGATTAAACCGTAGAGCCGCTAGCTGCGGAGCATTAATCTCAAAAGCAAGAAAT 7860
Db 7801 GCAAAAGAGATTAAACCGTAGAGCCGCTAGCTGCGGAGCATTAATCTCAAAAGCAAGAAAT 7860
Qy 7861 GAATGAAATTCAGAGGTTGAAATTTGAGAGAGATATTTCAATGAAAGAAATATAGATACT 7920
Db 7861 GAATGAAATTCAGAGGTTGAAATTTGAGAGAGATATTTCAATGAAAGAAATATAGATACT 7920
Qy 7921 AGCCCTTCTAAAGAAATTTGAGAGAGAAATCAATGTCAAAGTGGAAAAAGAAAAACAGAGTC 7980
Db 7921 AGCCCTTCTAAAGAAATTTGAGAGAGAAATCAATGTCAAAGTGGAAAAAGAAAAACAGAGTC 7980
Qy 7981 ACTGCTGAATCTCAAGAGAGCTTCTGTAGAGAGAGAGCAAGGGGCGAAATTTATCCGAA 8040
Db 7981 ACTGCTGAATCTCAAGAGAGCTTCTGTAGAGAGAGAGCAAGGGGCGAAATTTATCCGAA 8040
Qy 8041 GCAAAAGATGCGGAGAGCTTATTTGAAAGATTAATGAAATTCGGAAGAAAGTATTTT 8100
Db 8041 GCAAAAGATGCGGAGAGCTTATTTGAAAGATTAATGAAATTCGGAAGAAAGTATTTT 8100
Qy 8101 CATGCAAGATATGTAATATGAGAACCAACATTAATAATGAAGTAACAGCAGTTCTTAA 8160
Db 8101 CATGCAAGATATGTAATATGAGAACCAACATTAATAATGAAGTAACAGCAGTTCTTAA 8160
Qy 8161 GCAAGTAACAGTTCTGTATTTGGAGAGAGTTGAGTCAACCAAGGAGAACTCTGCTGCA 8220
Db 8161 GCAAGTAACAGTTCTGTATTTGGAGAGAGTTGAGTCAACCAAGGAGAACTCTGCTGCA 8220
Qy 8221 GGTAAAACTATGTAAGTAAGTTGAGAGAGAAATTTGTTCAAGAACAAATGATTTGATGCA 8280
Db 8221 GGTAAAACTATGTAAGTAAGTTGAGAGAGAAATTTGTTCAAGAACAAATGATTTGATGCA 8280
Qy 8281 ATTTCTAAAGTAAGAGTTTGGATGAGATTAAGTAACTGCTTAAATCTTCTGTAGATCA 8340
Db 8281 ATTTCTAAAGTAAGAGTTTGGATGAGATTAAGTAACTGCTTAAATCTTCTGTAGATCA 8340
Qy 8341 GGAATGAGAGAGAAATTTCCCGGAGCAGAGAGTGAATCTTCTACAGCAAAAGTAATACT 8400
Db 8341 GGAATGAGAGAGAAATTTCCCGGAGCAGAGAGTGAATCTTCTACAGCAAAAGTAATACT 8400
Qy 8401 GAATCCGTAAGTCTGTTTACGAAAGCAAGTTATGAAATAATGATTAACCAAAAAATAT 8460
Db 8401 GAATCCGTAAGTCTGTTTACGAAAGCAAGTTATGAAATAATGATTAACCAAAAAATAT 8460
Qy 8461 ATTTGAGAGAGTCAATGCTCTTCTTTAAATGATACAAAGAAATGAGCGAAATATAGATCT 8520
Db 8461 ATTTGAGAGAGTCAATGCTCTTCTTTAAATGATACAAAGAAATGAGCGAAATATAGATCT 8520

```

Qy 8521 TTAGCGTAGCCGGTGTGTCATGCAAGGAAACAAACAGCATTTACGAGATCAAAACAG 8580
Db 8521 TTAGCGTAGCCGGTGTGTCATGCAAGGAAACAAACAGCATTTACGAGATCAAAACAG 8580
Qy 8581 TTTAATCTTACACCTGTAAATGAGAGAAAAGTATCTCAACTTCGTGCAAAAGCTTTGGCT 8640
Db 8581 TTTAATCTTACACCTGTAAATGAGAGAAAAGTATCTCAACTTCGTGCAAAAGCTTTGGCT 8640
Qy 8641 AAAAATGAAAATTTATGAAAATGTAAGAGAACTGAGAGAGCTTTAGTGGAGCGGAAAACA 8700
Db 8641 AAAAATGAAAATTTATGAAAATGTAAGAGAACTGAGAGAGCTTTAGTGGAGCGGAAAACA 8700
Qy 8701 GCAGCCGTTGAAAATTTATACAAAGAGTACTACAGAGAGATTGGTTGCAAGAAATTTGGAA 8760
Db 8701 GCAGCCGTTGAAAATTTATACAAAGAGTACTACAGAGAGATTGGTTGCAAGAAATTTGGAA 8760
Qy 8761 ATTGAGATTAATTAATGAAAACGATTCGAAAGAGTAATAGATTGTAAGAGTCAACGAGAC 8820
Db 8761 ATTGAGATTAATTAATGAAAACGATTCGAAAGAGTAATAGATTGTAAGAGTCAACGAGAC 8820
Qy 8821 GGAACCAAGAGAGTCTTGTGCGAAAAGATGATATTTCTGTGAAAATTAACAATTTCAAGG 8880
Db 8821 GGAACCAAGAGAGTCTTGTGCGAAAAGATGATATTTCTGTGAAAATTAACAATTTCAAGG 8880
Qy 8881 GAAACAAAATCATTCATTGAAAGATTAAGCAGAAATTTGGAACCGGAAGTAAATGTA 8940
Db 8881 GAAACAAAATCATTCATTGAAAGATTAAGCAGAAATTTGGAACCGGAAGTAAATGTA 8940
Qy 8941 GATGCTTTGGAATGAACCTGATGATATCTACAGGAAAAGTGTGCTTATGTGGAATTT 9000
Db 8941 GATGCTTTGGAATGAACCTGATGATATCTACAGGAAAAGTGTGCTTATGTGGAATTT 9000
Qy 9001 GGTATTTGGAATTTGATGATTAATATGATGATTAAGAAAATGTAAGAACCCAAATTCGGA 9060
Db 9001 GGTATTTGGAATTTGATGATTAATATGATGATTAAGAAAATGTAAGAACCCAAATTCGGA 9060
Qy 9061 AGACATGCTATTTGTGAAAACCTACGAAAACAAAGAAATTCAGGATTTACAAGACAAA 9120
Db 9061 AGACATGCTATTTGTGAAAACCTACGAAAACAAAGAAATTCAGGATTTACAAGACAAA 9120
Qy 9121 GTAAATATTTCTTGGAAAAGAGACGCTGCACTGCACTGCAATTCGAATGTACACTT 9180
Db 9121 GTAAATATTTCTTGGAAAAGAGACGCTGCACTGCACTGCAATTCGAATGTACACTT 9180
Qy 9181 TCCAAATGAGATGATTAATTAATTTGGCAAGAGCTGCACTTCTCAATTAATTAAC 9240
Db 9181 TCCAAATGAGATGATTAATTAATTTGGCAAGAGCTGCACTTCTCAATTAATTAAC 9240
Qy 9241 AAAAATTCAAAAATTAATTAATTTAGCATCAAGTGAATCGAATGTGATGTTGAT 9300
Db 9241 AAAAATTCAAAAATTAATTAATTTAGCATCAAGTGAATCGAATGTGATGTTGAT 9300
Qy 9301 GGGGTGCTGAAAGCAAGAGCTGCAAGACCCAAAGGACAGTTAGTTAAGATCAATTA 9360
Db 9301 GGGGTGCTGAAAGCAAGAGCTGCAAGACCCAAAGGACAGTTAGTTAAGATCAATTA 9360
Qy 9361 AATAGAACTAATTAATGTTGATTTAGCAGAGAAAATTTAAAACAGAGGAAAACATCA 9420
Db 9361 AATAGAACTAATTAATGTTGATTTAGCAGAGAAAATTTAAAACAGAGGAAAACATCA 9420
Qy 9421 TATGCCGATATGATTAATTAATTAATTAAGTAAGACAAATTCGAAGCTATTTGGAT 9480
Db 9421 TATGCCGATATGATTAATTAATTAATTAAGTAAGTAAGACAAATTCGAAGCTATTTGGAT 9480
Qy 9481 GCCCAAAATCATGCTGCAAGCTGCTTGGCACTGCACTTGGAAAAATGAACTAAA 9540
Db 9481 GCCCAAAATCATGCTGCAAGCTGCTTGGCACTGCACTTGGAAAAATGAACTAAA 9540
Qy 9541 TTTAATTAATGAGATCCGAGAAATTTAAAATTAATCTGGCAAGATTTGGAGGGGAAAGCTA 9600
Db 9541 TTTAATTAATGAGATCCGAGAAATTTAAAATTAATCTGGCAAGATTTGGAGGGGAAAGCTA 9600
Qy 9601 AAAAAACGTGCGTAGATCTTAATCAGGTAGACTGATTAAGGATTAATATACATGCGAT 9660

```

```

Db 9601 AAAAAACGTGCGTAGATCTTAATCAGGTAGACTGATTAAGGATTAATATACATGCGAT 9660
Qy 9661 TCTTCTGAAAAAGCATACAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGGAAA 9720
Db 9661 TCTTCTGAAAAAGCATACAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGGAAA 9720
Qy 9721 AAATGA 9726
Db 9721 AAATGA 9726

RESULT 2
AA167647
ID AA167647 standard; DNA; 11130 BP.
XX
XX AA167647;
AC 27-FEB-2002 (first entry)
XX
XX F. necrophorum leukotoxin gene sequence.
DE
XX Leukotoxin; infection; immunisation; liver abscess; foot rot;
XX bactericidal; vaccine; ds.
XX F. necrophorum necrophorum.
XX WO200180886-A2.
XX
XX 01-NOV-2001.
PD
XX 25-APR-2001; 2001WO-US013240.
PF
XX 25-APR-2000; 2000US-00558257.
PR 24-APR-2001; 2001US-00841786.
XX
XX (UNIV ) UNIV KANSAS STATE RES FOUND.
PA
XX Nagareja TG, Stewart GC, Narayanan SK, Chengappa MM,
PI WPI; 2002-049245/06.
XX
XX F. necrophorum necrophorum polypeptide useful as vaccine in immunizing an
PT animal against an infection e.g. foot rot, or liver abscesses caused by
PT the bacterium.
XX
XX Disclosure; Page 102-107; 108pp; English.
XX
XX The invention provides an isolated F. necrophorum necrophorum leukotoxin
CC polypeptide and its truncated versions. The leukotoxin protein can be
CC expressed by standard recombinant methodology. The leukotoxin gene is
CC useful for preparing a vaccine which confers effective immunity against
CC infection caused by F. necrophorum. The vaccine is useful for immunising
CC an animal against liver abscesses caused by F. necrophorum and for
CC preventing foot rot caused by F. necrophorum infection. The present
CC sequence represents the F. necrophorum full-length leukotoxin gene
CC sequence
XX
XX Sequence 11130 BP; 4205 A; 1472 C; 2511 G; 2942 T; 0 U; 0 Other;
SQ
Query Match 99.9%; Score 9718; DB 6; Length 11130;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 9721; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 ATGAGCGGCATCAAAAATTAACGTTCAAGAGCAAGAGAGAGATTCAGATTTCAAAAA 60
Db 1034 ATGAGCGGCATCAAAAATTAACGTTCAAGAGCAAGAGAGAGATTCAGATTTCAAAAA 1093
Qy 61 GTTTAATGATTTGGATTTGTTATTAACACTATGACCGTGAAGGCTTAATGATCAATC 120
Db 1094 GTTTAATGATTTGGATTTGTTATTAACACTATGACCGTGAAGGCTTAATGATCAATC 1153
Qy 121 ACCGCACTGAGATTTTGAACAAAAATGAAAAAAGATTAATGTTATGACATTA 180

```

1154 ACCGGACCTGAGATTTTGGAACTAAATAGAAAAAAGATATGTTTATGACATTTACT 1213
181 ACAAACAAGATTCAAGGGGAGAACGCTTTTAAACGTTTATATAGATTGGCTTTAACAGAA 240
1214 ACAAACAAGATTCAAGGGGAGAACGCTTTTAAACGTTTATATAGATTGGCTTTAACAGAA 1273
241 AATAATATGCAAAATCTATATTTTGGGAAAAAGATAGACGGGGTAAATATCTTTT 300
1274 AATAATATGCAAAATCTATATTTTGGGAAAAAGATAGACGGGGTAAATATCTTTT 1333
301 AACTTTGTCATATGAAAAATTTGAAGATATGAGATTATCAACGAAATTCAGAAAAATAA 360
1334 AACTTTGTCATATGAAAAATTTGAAGATATGAGATTATCAACGAAATTCAGAAAAATAA 1393
361 ATTGGAGAAATTTATATTTCTTAAGCTCGAAAGGATGCAATGAGAAATATGAGATT 420
1394 ATTGGAGAAATTTATATTTCTTAAGCTCGAAAGGATGCAATGAGAAATATGAGATT 1453
421 ATCAAATGCTGTTCTTTTCAATTCATATTTCCAAAACAAGATTTTAAAGAGCTTTG 480
1454 ATCAAATGCTGTTCTTTTCAATTCATATTTCCAAAACAAGATTTTAAAGAGCTTTG 1513
481 GAAGAAAGCCAAACATGCTAAAGTTTAAATGAAATGATCCAGTAGATGAAAGATGAAA 540
1514 GAAGAAAGCCAAACATGCTAAAGTTTAAATGAAATGATCCAGTAGATGAAAGATGAAA 1573
541 ATTCATTTGAAATCCGAAATGAAAGCATTAAGTGAAGAAAGAAATCAATCTGTTGAAGC 600
1574 ATTCATTTGAAATCCGAAATGAAAGCATTAAGTGAAGAAAGAAATCAATCTGTTGAAGC 1633
601 ATCGGTTTATATGCGGCGGATTTAGATTGAAGATATGCAATCTTAAGACAGAAAT 660
1634 ATCGGTTTATATGCGGCGGATTTAGATTGAAGATATGCAATCTTAAGACAGAAAT 1693
661 AACAATTTTAAATTTTAGCAATATAGATGCAATTAATCTGCTGACCGGAGAT 720
1694 AACAATTTTAAATTTTAGCAATATAGATGCAATTAATCTGCTGACCGGAGAT 1753
721 TTAAAGCTACCAAGACAAATCTGAGATATATTTCTTCAAGCTCACAATGATTTCTCT 780
1754 TTAAAGCTACCAAGACAAATCTGAGATATATTTCTTCAAGCTCACAATGATTTCTCT 1813
781 CAAAAAGCTATGCGAAAAATTTCACTGTTGAAAGAAATAGAAATATGTTAAAGCA 840
1814 CAAAAAGCTATGCGAAAAATTTCACTGTTGAAAGAAATAGAAATATGTTAAAGCA 1873
841 AATACCAAGCAAAATTTGAAATCTGATCTGATTTGGAAGCAGATGGAATATTAATAAT 900
1874 AATACCAAGCAAAATTTGAAATCTGATCTGATTTGGAAGCAGATGGAATATTAATAAT 1933
901 AGTGCAGAAAGCTACAATGAGGAGATTTATTAAGAAAGAGGAAAAAGAACTTATTAAC 960
1934 AGTGCAGAAAGCTACAATGAGGAGATTTATTAAGAAAGAGGAAAAAGAACTTATTAAC 1993
961 ACTCTTTTAAAGTTTATCAGATGAGAGCTTCCGTAAAGATTAATAAGAAAAAGTCATA 1020
1994 ACTCTTTTAAAGTTTATCAGATGAGAGCTTCCGTAAAGATTAATAAGAAAAAGTCATA 2053
1021 GGAAGAAATTTGACATTAACGCTGAAGCAAGAAATTTCTATGATGCAATTTTATGTTACT 1080
2054 GGAAGAAATTTGACATTAACGCTGAAGCAAGAAATTTCTATGATGCAATTTTATGTTACT 2113
1081 AAGCTTGCAAGACCTTTTATGCTTGTATACAGGTTCTATTTCTCTATCAATTTTAAT 1140
2114 AAGCTTGCAAGACCTTTTATGCTTGTATACAGGTTCTATTTCTCTATCAATTTTAAT 2173
1141 GGAATTTTAAAGTTTATGCAAGTATGATGCTGTTATGGAAGAAAGATGCCAAAGTC 1200
2174 GGAATTTTAAAGTTTATGCAAGTATGATGCTGTTATGGAAGAAAGATGCCAAAGTC 2233
1201 GAAGCAACAGAAAGGCAAAATATCTTCTTACGCTGGAATGAAGACATATGCGA 1260

2234 GAAGCAACAGAAAGGCAAAATATCTTCTTACATGAGATGAAGCAATATGCGA 2293
1261 GCACTTACTTCTCCATTTAAAAATTTACCAATTTATTTGGAAGAACCAATGGAATCTT 1320
2294 GCACTTACTTCTCCATTTAAAAATTTACCAATTTATTTGGAAGAACCAATGGAATCTT 2353
1321 CTCACTATCGAGCGGATATATTTCTGCAAAAAGTAAATTCCAATGTAATATGGAAGA 1380
2354 CTCACTATCGAGCGGATATATTTCTGCAAAAAGTAAATTCCAATGTAATATGGAAGA 2413
1381 GAAGTAAATTCGAAGGAGAAAGCAGATTTACTTCAAAATCTGAATAATCTATGATGCT 1440
2414 GAAGTAAATTCGAAGGAGAAAGCAGATTTACTTCAAAATCTGAATAATCTATGATGCT 2473
1441 TCTGTTTCTGTTGGAACGATGAGAAATTTCCAAATGAGCTTCTGATATTTGGAGC 1500
2474 TCTGTTTCTGTTGGAACGATGAGAAATTTCCAAATGAGCTTCTGATATTTGGAGC 2533
1501 GAAGGAGAAATTAATTTCTCCGTCAAGATTTGCTAAAGGACAAAGTAAATCAGAAACG 1560
2534 GAAGGAGAAATTAATTTCTCCGTCAAGATTTGCTAAAGGACAAAGTAAATCAGAAACG 2593
1561 GATGATGTAATGTAAGAAAGTGAAGCAGATTAATTCATTCGAGCTGCTGTAAGAGTGA 1620
2594 GATGATGTAATGTAAGAAAGTGAAGCAGATTAATTCATTCGAGCTGCTGTAAGAGTGA 2653
1621 TTGGGGATATGCTGTAATGCGGTTGCTGCAATATTTCTTAATATGCTTCTCTC 1680
2654 TTGGGGATATGCTGTAATGCGGTTGCTGCAATATTTCTTAATATGCTTCTCTC 2713
1681 CGTATGATGTAAGTGAATGATATCTACATGCAAGAAAGCAGTAAATGTAAGGCTCATAC 1740
2714 CGTATGATGTAAGTGAATGATATCTACATGCAAGAAAGCAGTAAATGTAAGGCTCATAC 2773
1741 ATTAATTAATAATATGATTTCTGCAAAACAGATCTGATTTGGAACTTCCAAATTTATGAA 1800
2774 ATTAATTAATAATATGATTTCTGCAAAACAGATCTGATTTGGAACTTCCAAATTTATGAA 2833
1801 GATCAGTTTATGAAATCAGGTCATCTAAATCAATTTTATGAGCAATTAACAGCGGTTT 1860
2834 GATCAGTTTATGAAATCAGGTCATCTAAATCAATTTTATGAGCAATTAACAGCGGTTT 2893
1861 GGAGGAGACAGTGTCAATGAGGAAATTAAGAAATTAAGCTAACGAACTTATGATGTCGCT 1920
2894 GGAGGAGACAGTGTCAATGAGGAAATTAAGAAATTAAGCTAACGAACTTATGATGTCGCT 2953
1921 GTGCTGCAACCATATGCAATCATATTAATTTCTGCTTCTGCTGCAATGAGAGAGTGA 1980
2954 GTGCTGCAACCATATGCAATCATATTAATTTCTGCTTCTGCTGCAATGAGAGAGTGA 3013
1981 AGACTTTCTTCAAGAGTGAAGGAGATTAATGTAAGGCAATTAATGAGCTCAAAATCTT 2040
3014 AGACTTTCTTCAAGAGTGAAGGAGATTAATGTAAGGCAATTAATGAGCTCAAAATCTT 3073
2041 CGAGGACATACGTCAGAGTGAAGGAGTGTGCTGTAAGGAGGAGGAAAAAGAACTTAT 2100
3074 CGAGGACATACGTCAGAGTGAAGGAGTGTGCTGTAAGGAGGAGGAAAAAGAACTTAT 3133
2101 GGAATATGACAGCTTTTATGGAACATTAATAATATGCTTCTGTAACAATGCGAT 2160
3134 GGAATATGACAGCTTTTATGGAACATTAATAATATGCTTCTGTAACAATGCGAT 3193
2161 CATGCTGAATGCTATCGAGAGGAAATTTGATATCAACAGTGAATAATAATTTGAATAT 2220
3194 CATGCTGAATGCTATCGAGAGGAAATTTGATATCAACAGTGAATAATAATTTGAATAT 3253
2221 AAAAATCTCTCAAAAATGCGAAAGCTGTTATGATTAATATAGAACTTTTAAAGAGCT 2280
3254 AAAAATCTCTCAAAAATGCGAAAGCTGTTATGATTAATATAGAACTTTTAAAGAGCT 3313
2281 TTGGAAGAAAGAAAGCAAACTCCAGAAATATGATCCGAAGATATGAAATCTATTTGAAAA 2340
3314 TTGGAAGAAAGAAAGCAAACTCCAGAAATATGATCCGAAGATATGAAATCTATTTGAAAA 3373

QY 2341 TTATTGAATGCATTTTCAGAAAAATTGATGAAAAACCGAGCTTTTACTAAATGCTGAA 2400
DB 3374 TTATTGAATGCATTTTCAGAAAAATTGATGAAAAACCGAGCTTTTACTAAATGCTGAA 3433
QY 2401 AGAATGACAAATTAATCTTCCCGATGGAATCTCAAAAAACGAAACGCTCTATGAATTTGCA 2460
DB 3434 AGAATGACAAATTAATCTTCCCGATGGAATCTCAAAAAACGAAACGCTCTATGAATTTGCA 3493
QY 2461 AACTATGTTACAGGAGAAATGAAAAATTAGAGAAAAATTACCGAAAGATTTTAAAGCT 2520
DB 3494 AACTATGTTACAGGAGAAATGAAAAATTAGAGAAAAATTACCGAAAGATTTTAAAGCT 3553
QY 2521 TTTTCAGAAAGATTGATGAGTGAATTAAGAACTTTGAATTTTACAGAGATAGAAAT 2580
DB 3554 TTTTCAGAAAGATTGATGAGTGAATTAAGAACTTTGAATTTTACAGAGATAGAAAT 3613
QY 2581 TATGCAATTTTTCACACTTTTACCTCTCCGAGCTAATGAGAAAGAGATTTCTCT 2640
DB 3614 TATGCAATTTTTCACACTTTTACCTCTCCGAGCTAATGAGAAAGAGATTTCTCT 3673
QY 2641 GTGGAGAGAGCTGTTTCGTGGTGAACAAGAAATTATGCAAGGTAACGTTGGAATA 2700
DB 3674 GTGGAGAGAGCTGTTTCGTGGTGAACAAGAAATTATGCAAGGTAACGTTGGAATA 3733
QY 2701 GAGAGTAAACTTGCTGCAAAAAAAGATTTAAATATPAAAAGCTATCAATTAAGCAGAAACA 2760
DB 3734 GAGAGTAAACTTGCTGCAAAAAAAGATTTAAATATPAAAAGCTATCAATTAAGCAGAAACA 3793
QY 2761 GTGAATTTAGTTGGAAATTTTGACCTTGGAGAAACGATCAATCCGAAAGTGCATCGGA 2820
DB 3794 GTGAATTTAGTTGGAAATTTTGACCTTGGAGAAACGATCAATCCGAAAGTGCATCGGA 3853
QY 2821 GGAAGATTAAATGTTCAAAAGATCGAAAAATTCAGCTATGTAAGAGCTPAAAAGAAAGCT 2880
DB 3854 GGAAGATTAAATGTTCAAAAGATCGAAAAATTCAGCTATGTAAGAGCTPAAAAGAAAGCT 3913
QY 2881 GAATTAATCAGAGAAAAATTTAATGCAATGCAATGGAACAGACTTTTCATGTAGCGGA 2940
DB 3914 GAATTAATCAGAGAAAAATTTAATGCAATGCAATGGAACAGACTTTTCATGTAGCGGA 3973
QY 2941 TCTTTTAATGTGCTCAGCTGGAATGCAATCAATGGAATGGAAGTTATAGTGAAGT 3000
DB 3974 TCTTTTAATGTGCTCAGCTGGAATGCAATCAATGGAATGGAAGTTATAGTGAAGT 4033
QY 3001 ATCAGTAAGGCAAGAGTTCCATGTAGCAAGCAATATTGAAAGCTPATAAAAAATTT 3060
DB 4034 ATCAGTAAGGCAAGAGTTCCATGTAGCAAGCAATATTGAAAGCTPATAAAAAATTT 4093
QY 3061 GCTTTAAACAGTAAGATGATCTTCTGTTTGAATGCTGCCGTTTCAGCGGGAATCGGA 3120
DB 4094 GCTTTAAACAGTAAGATGATCTTCTGTTTGAATGCTGCCGTTTCAGCGGGAATCGGA 4153
QY 3121 ACGAAAAATGCGCGCGGTGCGGGTGTCTGTCGCGTAATGATTAATTTCAACAA 3180
DB 4154 ACGAAAAATGCGCGCGGTGCGGGTGTCTGTCGCGTAATGATTAATTTCAACAA 4213
QY 3181 GCTTCATTAAGATTAATGACGAAGGACAAAGTAATATGATTAAGATTAAGATGAA 3240
DB 4214 GCTTCATTAAGATTAATGACGAAGGACAAAGTAATATGATTAAGATTAAGATGAA 4273
QY 3241 GTAAAGTAATCTGCGGAATCTTTAGAGATGATGCAAAAAACGACGGAACAATCAACGT 3300
DB 4274 GTAAAGTAATCTGCGGAATCTTTAGAGATGATGCAAAAAACGACGGAACAATCAACGT 4333
QY 3301 ATTTCTGTTGCGGAGGAATTAATAAGTTGGAAGTAAACGAGTGAAGAAAAACCGAA 3360
DB 4334 ATTTCTGTTGCGGAGGAATTAATAAGTTGGAAGTAAACGAGTGAAGAAAAACCGAA 4393
QY 3361 TCAGAGAAAGACGAGAGGATTTTGTGCAAAATTCGGAACAAGTGAAGCTCTGTAATA 3420
DB 4394 TCAGAGAAAGACGAGAGGATTTTGTGCAAAATTCGGAACAAGTGAAGCTCTGTAATA 4453

QY 3421 AATAAAATTCAGATAGTATGATTCATTAACAGAAAAATTCAATTAATTAATTCGAA 3480
DB 4454 AATAAAATTCAGATAGTATGATTCATTAACAGAAAAATTCAATTAATTAATTCGAA 4513
QY 3481 GGAATPAAAAAAGCGGGGAATCTTCCTTCGAACGTTTCTCATCTCCGATPAAAGAACG 3540
DB 4514 GGAATPAAAAAAGCGGGGAATCTTCCTTCGAACGTTTCTCATCTCCGATPAAAGAACG 4573
QY 3541 TCTTTCAGTTGGAGCTTCTCGAAAGTGTTCCTTCAATTAATTAATAAAGAAACATCT 3600
DB 4574 TCTTTCAGTTGGAGCTTCTCGAAAGTGTTCCTTCAATTAATTAATAAAGAAACATCT 4633
QY 3601 GCTGCTGATAGTGAATGAATTAATTAAGAGGACAAATPAAAAGTGAAGTGAATCT 3660
DB 4634 GCTGCTGATAGTGAATGAATTAATTAAGAGGACAAATPAAAAGTGAAGTGAATCT 4693
QY 3661 TCTTCTGATTTACTTTTGTGGAGCATGGGGCCGATCTGCTGCACTTCAATGGAATCAT 3720
DB 4694 TCTTCTGATTTACTTTTGTGGAGCATGGGGCCGATCTGCTGCACTTCAATGGAATCAT 4753
QY 3721 ATTGGAATGGAATATAGCAATCATGCTGCTGTTAGCTGAGACGCTGCTGTAATTAAT 3780
DB 4754 ATTGGAATGGAATATAGCAATCATGCTGCTGTTAGCTGAGACGCTGCTGTAATTAAT 4813
QY 3781 ATTCAAAGTAAACAAAGTCTTGTGTTAAAAATAGTGAATTCGAATGCCAATAATTT 3840
DB 4814 ATTCAAAGTAAACAAAGTCTTGTGTTAAAAATAGTGAATTCGAATGCCAATAATTT 4873
QY 3841 AAAGTAAATGCTTTGAGTGAAGGAATCAATAGAGAGAGAGAGAGGTTTGGAGCAATT 3900
DB 4874 AAAGTAAATGCTTTGAGTGAAGGAATCAATAGAGAGAGAGAGAGGTTTGGAGCAATT 4933
QY 3901 AAAGAAATGGAAGCAAGAAAAAATATGTAATTTGGAACCTTCTCTATCAACTTA 3960
DB 4934 AAAGAAATGGAAGCAAGAAAAAATATGTAATTTGGAACCTTCTCTATCAACTTA 4993
QY 3961 GTGAACAAATGAAGTTTCTGCAAAATCAGAAAAATTAATCAGTAAGCAGAGAAATCTGAAAGC 4020
DB 4994 GTGAACAAATGAAGTTTCTGCAAAATCAGAAAAATTAATCAGTAAGCAGAGAAATCTGAAAGC 5053
QY 4021 CAAAAAATGGAATGTTGATGTCATGCTCTTATCAAGCGGACACCCAAATGACAGAGCTTTA 4080
DB 5054 CAAAAAATGGAATGTTGATGTCATGCTCTTATCAAGCGGACACCCAAATGACAGAGCTTTA 5113
QY 4081 AATTTCAAGCTGGAAGATCAAAATGGAATCTGTAAGGAGCTACTGTGACCAATTA 4140
DB 5114 AATTTCAAGCTGGAAGATCAAAATGGAATCTGTAAGGAGCTACTGTGACCAATTA 5173
QY 4141 AACAACAAAGTAATGCTTCTATTAGTGTGGAGATATTAACGTTAATCGACCGAC 4200
DB 5174 AACAACAAAGTAATGCTTCTATTAGTGTGGAGATATTAACGTTAATCGACCGAC 5233
QY 4201 GCAAAAGCTCTTTTATGCAACCATCTCAATGATCTGTGCAAGTGAACGAGAGGACAAATT 4260
DB 5234 GCAAAAGCTCTTTTATGCAACCATCTCAATGATCTGTGCAAGTGAACGAGAGGACAAATT 5293
QY 4261 AGTTCTGAGCGGGAATTAAGAAATTAACAAGGCGCTGTTCTGTCAATTAAGTGAAT 4320
DB 5294 AGTTCTGAGCGGGAATTAAGAAATTAACAAGGCGCTGTTCTGTCAATTAAGTGAAT 5353
QY 4321 GACGTGAAGGTAGCGTTGATTAATCTTCAATCGAAGGAGCTTAATGAATCAATGTCATT 4380
DB 5354 GACGTGAAGGTAGCGTTGATTAATCTTCAATCGAAGGAGCTTAATGAATCAATGTCATT 5413
QY 4381 GCCAAAGTGTCAAAAGGAAGTTCTGATTAAGCAAAAGAAATATCAAGCTTTACTAAATGGA 4440
DB 5414 GCCAAAGTGTCAAAAGGAAGTTCTGATTAAGCAAAAGAAATATCAAGCTTTACTAAATGGA 5473
QY 4441 AAAGATTAATAAATTTTGAAGATGCTGCTATTAATAGCATGGAATGTTATTAATACG 4500
DB 5474 AAAGATTAATAAATTTTGAAGATGCTGCTATTAATAGCATGGAATGTTATTAATACG 5533
QY 4501 AAAGAAACAATTAAGAAAAAGCAAAAGAAAAAGAGAGCGGTCATTGTAATGCTGCTTTA 4560

Db 5534 AAGGACAACTAGAAAAAGCAAAAGAAAAAGAGCGGTCTATTGTAAATCTCTTTA 5593
Qy 4561 TCGGTTGCTGGAACGGATTAATCCGCTGAGAGTAGCTATTGTCAGTCATACTGTAA 4620
Db 5594 TCGGTTGCTGGAACGGATTAATCCGCTGAGAGTAGCTATTGTCAGTCATACTGTAA 5653
Qy 4621 AATAAATTTAAAGCAGATTGATGAGAACATTAAGAAAGCCGAGAGATTAATTCAT 4680
Db 5654 AATAAATTTAAAGCAGATTGATGAGAACATTAAGAAAGCCGAGAGATTAATTCAT 5713
Qy 4681 GCGAAACATGTAAATGTGAGGCAAAATCATCTACTGTGTGTGTAATCGGCTCTGGA 4740
Db 5714 GCGAAACATGTAAATGTGAGGCAAAATCATCTACTGTGTGTGTAATCGGCTCTGGA 5773
Qy 4741 CTTCCTATCAGCAAGATGCTTTTTCAGGAATGCGATCTGAGCATGCGAAGATTATCA 4800
Db 5774 CTTCCTATCAGCAAGATGCTTTTTCAGGAATGCGATCTGAGCATGCGAAGATTATCA 5833
Qy 4801 AATGACAGATTGCAAAAGGTGATTAAGAAAGAAATTTCTGCTGATTCTTTAAATGTGAAC 4860
Db 5834 AATGACAGATTGCAAAAGGTGATTAAGAAAGAAATTTCTGCTGATTCTTTAAATGTGAAC 5893
Qy 4861 GGAATTAATTCATCTTGGGGTGAAATGTTGGGGGAACCATGGCGGTTCTCTTTCTAG 4920
Db 5894 GGAATTAATTCATCTTGGGGTGAAATGTTGGGGGAACCATGGCGGTTCTCTTTCTAG 5953
Qy 4921 GCGGTAGAGCTGCTTTTTCGGAATATCTCTCATATATAAACTCTGCTTGATTACA 4980
Db 5954 GCGGTAGAGCTGCTTTTTCGGAATATCTCTCATATATAAACTCTGCTTGATTACA 6013
Qy 4981 GGAACGAGGTAAATCTTTTATGTGAAAGAAATCAAAAGTCATGTACAGCTTTGAAT 5040
Db 6014 GGAACGAGGTAAATCTTTTATGTGAAAGAAATCAAAAGTCATGTACAGCTTTGAAT 6073
Qy 5041 GATTCTCATATTAACAAACGTTTCTGCTGAGAGCGCTGCAAGTATTAAGAGCGTGAATC 5100
Db 6074 GATTCTCATATTAACAAACGTTTCTGCTGAGAGCGCTGCAAGTATTAAGAGCGTGAATC 6133
Qy 5101 GGAGAAATGATCTCTCATCGTGTCTGATGAACGAAAGCTTTAGTATGATTC 5160
Db 6134 GGAGAAATGATCTCTCATCGTGTCTGATGAACGAAAGCTTTAGTATGATTC 6193
Qy 5161 GAGTTGGAAGAGTAAATCTTTTCAATGTATGATGCAAAAGATCAAAATTAATACA 5220
Db 6194 GAGTTGGAAGAGTAAATCTTTTCAATGTATGATGCAAAAGATCAAAATTAATACA 6253
Qy 5221 ATTGCGGGAATTCGAATGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACA 5280
Db 6254 ATTGCGGGAATTCGAATGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACA 6313
Qy 5281 AATATTGAAAAAACAATCAGTTATAGCTATTGTAAAAAACAATAAATTACACGCGAAT 5340
Db 6314 AATATTGAAAAAACAATCAGTTATAGCTATTGTAAAAAACAATAAATTACACGCGAAT 6373
Qy 5341 GATCAAGATGAAAAAATATCAATGTGACTGCAAAAAGATTATACATGACCAATACTATA 5400
Db 6374 GATCAAGATGAAAAAATATCAATGTGACTGCAAAAAGATTATACATGACCAATACTATA 6433
Qy 5401 GCAGTGCGAGTTGAGAGACAAAGAGCCTCTGTGCAAGAGACTTCTGCAATACTACC 5460
Db 6434 GCAGTGCGAGTTGAGAGACAAAGAGCCTCTGTGCAAGAGACTTCTGCAATACTACC 6493
Qy 5461 TTGAATTAAGACAGTTTCTCTCATGTGATCAATCTGATATTTGACAAAGATTTAGAGAA 5520
Db 6494 TTGAATTAAGACAGTTTCTCTCATGTGATCAATCTGATATTTGACAAAGATTTAGAGAA 6553
Qy 5521 GAAAAATATGAAAAATAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATACGAGTCAA 5580
Db 6554 GAAAAATATGAAAAATAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATACGAGTCAA 6613
Qy 5581 GTGTGCAAAATGCGAAGTGTCTTCCGAGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCA 5640

Db 6614 GTGTGCAAAATGCGAAGTGTCTTCCGAGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCA 6673
Qy 5641 GTAGACGTTAATTAATAATTAACAAATATCTTGCACATATTAATAATAGTACTCAAAAT 5700
Db 6674 GTAGACGTTAATTAATAATTAACAAATATCTTGCACATATTAATAATAGTACTCAAAAT 6733
Qy 5701 GTACGAAATGTCTTGGTAAAAAGCAAAATCTCATCTATTTAAACAATTGGAATTGGA 5760
Db 6734 GTACGAAATGTCTTGGTAAAAAGCAAAATCTCATCTATTTAAACAATTGGAATTGGA 6793
Qy 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGTACAGGTTCTGTAGCAATGAATTAAGATTGTA 5820
Db 6794 GCTGAGTTGAGAGCTGAGAGAGCTGAGTACAGGTTCTGTAGCAATGAATTAAGATTGTA 6853
Qy 5821 AATAATAGATAGCAATTTAAATGATCAAAATCACTGCGCAAGGAAATGTGCGAGTT 5880
Db 6854 AATAATAGATAGCAATTTAAATGATCAAAATCACTGCGCAAGGAAATGTGCGAGTT 6913
Qy 5881 ATTACAGAGTCTGATGCGGTAAATTTGTAATTAATGAGAAACAGTCTGAGTGCCTGT 5940
Db 6914 ATTACAGAGTCTGATGCGGTAAATTTGTAATTAATGAGAAACAGTCTGAGTGCCTGT 6973
Qy 5941 GCAGCAATAGAGCCTCAACAGTGTGAATGAATTAACAGATCTACAAACCATATGTA 6000
Db 6974 GCAGCAATAGAGCCTCAACAGTGTGAATGAATTAACAGATCTACAAACCATATGTA 7033
Qy 6001 AAGATTTCAAGTATGCTTAAAGAAACAGATGATTAATTAATCTGCAAGGCGAA 6060
Db 7034 AAGATTTCAAGTATGCTTAAAGAAACAGATGATTAATTAATCTGCAAGGCGAA 7093
Qy 6061 GTTGAATTAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6120
Db 7094 GTTGAATTAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7153
Qy 6121 AAAAAAATAATAGTAATTAATAAAGATTTGTTACCAATAGTTACGTACTACTACTTA 7213
Db 7154 AAAAAAATAATAGTAATTAATAAAGATTTGTTACCAATAGTTACGTACTACTACTTA 7213
Qy 6181 AATCTTTATTTGGCAATATCCCTGTGTACAGCAACGCGAGTGTGCAAGATCTGTAAT 6240
Db 7214 AATCTTTATTTGGCAATATCCCTGTGTACAGCAACGCGAGTGTGCAAGATCTGTAAT 7273
Qy 6241 ATCAACAAGTTATGAGAGAAACAGAGCTTTGTGAAATCTTATTTAAATGCAAAA 6300
Db 7274 ATCAACAAGTTATGAGAGAAACAGAGCTTTGTGAAATCTTATTTAAATGCAAAA 7333
Qy 7334 CATTAATCTGTAATAATCAGAGATTACAGAAATTCATCGAGTAGTAAAGTTCTGTGAT 7393
Db 6301 CATTAATCTGTAATAATCAGAGATTACAGAAATTCATCGAGTAGTAAAGTTCTGTGAT 7453
Qy 6361 GTTGTGGAATGTAGAGTAGAGGCTTCTTGTGATCCAAATTTTAAAAAGAAATACC 6420
Db 7394 GTTGTGGAATGTAGAGTAGAGGCTTCTTGTGATCCAAATTTTAAAAAGAAATACC 7453
Qy 6421 AAGACAAGAGTTGGAATACTCAATGTGATGAAGTTTGGAGAAAGAGCTGAATT 6480
Db 7454 AAGACAAGAGTTGGAATACTCAATGTGATGAAGTTTGGAGAAAGAGCTGAATT 7513
Qy 6481 ACAGCAGATTTTAAGCAAGAAATTTCTCTTTTGAAGTGTGCAAGCAGCGGGTA 6540
Db 7514 ACAGCAGATTTTAAGCAAGAAATTTCTCTTTTGAAGTGTGCAAGCAGCGGGTA 7573
Qy 6541 GAGCCGAGAGTGTGAGAGAACGTTTCCGTAAATCAATTTGAGAGAAAGAGAGTATGAT 6600
Db 7574 GAGCCGAGAGTGTGAGAGAACGTTTCCGTAAATCAATTTGAGAGAAAGAGAGTATGAT 7633
Qy 6601 GTGGAAGAGCAAAAGTTTTGTAAAAAAGCTGAGATTACAGCAAAAGCTTAATGTTCT 6660
Db 7634 GTGGAAGAGCAAAAGTTTTGTAAAAAAGCTGAGATTACAGCAAAAGCTTAATGTTCT 7693
Qy 6661 GTTGCAATTTGGAATGCGCAGTGTGAGTGTGCTGCAAAAGAGAGTGTGAGAGCA 6720
Db 7694 GTTGCAATTTGGAATGCGCAGTGTGAGTGTGCTGCAAAAGAGAGTGTGAGAGCA 7753

QY 6721 GTGGAGTTACCAAGATGATCAAAACAGAGAGCAGAGTGAATAATTTCTAAATTTATG 6780
DB 7754 GTGGCAGTTACCAAGATGATCAAAACAGAGAGCAGAGTGAATAATTTCTAAATTTATG 7813
QY 6781 ACTCGAACAAGTTAGATGATTAAGCAGAAAAAGATGAATAAATCAAGTAAGTCTGAAATCGGT 6840
DB 7814 ACTCGAACAAGTTAGATGATTAAGCAGAAAAAGATGAATAAATCAAGTAAGTCTGAAATCGGT 7873
QY 6841 TCAGCCGAGAGTGAATTTCTGACGCCGAGTATCTGAGTGGTCTTCTGCAATAATTT 6900
DB 7874 TCAGCCGAGAGTGAATTTCTGACGCCGAGTATCTGAGTGGTCTTCTGCAATAATTT 7933
QY 6901 GCAAAATAGGTAGAAACAGATATCGATCACTTACCTTCTCTCTCTGATGTAAT 6960
DB 7934 GCAAAATAGGTAGAAACAGATATCGATCACTTACCTTCTCTCTCTGATGTAAT 7993
QY 6961 GTAAAAAGCTCTTAATTAATAATTTTCGAAATTCCTTGAACAGCCGCTGAGAGAGCCGAGGTCTT 7020
DB 7994 GTAAAAAGCTCTTAATTAATAATTTTCGAAATTCCTTGAACAGCCGCTGAGAGAGCCGAGGTCTT 8053
QY 7021 GCAGAGTTACCGAGTGGTCTTCTGTTAACCTATTAATAGTTCTGTATAGCTGAGTT 7080
DB 8054 GCAGAGTTACCGAGTGGTCTTCTGTTAACCTATTAATAGTTCTGTATAGCTGAGTT 8113
QY 7081 CACAATTAATCTGATTTGACTTCGTTACGAGAAAAAGTAAATGTAAACGCAAAAGAGAA 7140
DB 8114 CACAATTAATCTGATTTGACTTCGTTACGAGAAAAAGTAAATGTAAACGCAAAAGAGAA 8173
QY 7141 AAAAAATTAATAGCAAAACAGCAGCAAAATGCGAAATCGAGAGCAAGCAATCGGAGCCAT 7200
DB 8174 AAAAAATTAATAGCAAAACAGCAGCAAAATGCGAAATCGAGAGCAAGCAATCGGAGCCAT 8233
QY 7201 GCTTGTTAATTAATTTTGGAAACAGCTGTAGAAAGATAGAAAAATTTGAAAGAAAAAGA 7260
DB 8234 GCTTGTTAATTAATTTTGGAAACAGCTGTAGAAAGATAGAAAAATTTGAAAGAAAAAGA 8293
QY 7261 ACAGAGTTTTTAAAACTTTAGACGAAAGTTAAACAAGAACAGATTAATAATGAT 7320
DB 8294 ACAGAGTTTTTAAAACTTTAGACGAAAGTTAAACAAGAACAGATTAATAATGAT 8353
QY 7321 GCTACGAAAAAAATCTTACATCAGAGGTAATTTCTACAGAAAGTATCTTCTGTAAGAGCG 7380
DB 8354 GCTACGAAAAAAATCTTACATCAGAGGTAATTTCTACAGAAAGTATCTTCTGTAAGAGCG 8413
QY 7381 GATAGAGGATTAATCAAGGAGAAAGAAATTAAGCATTTGTAAGATCTTCTGATATTAAT 7440
DB 8414 GATAGAGGATTAATCAAGGAGAAAGAAATTAAGCATTTGTAAGATCTTCTGATATTAAT 8473
QY 7441 GAAAAAATGTAGATATTAACAACAGAGCAAGAAATTAATATCATCTTCTACTGAGTTTG 7500
DB 8474 GAAAAAATGTAGATATTAACAACAGAGCAAGAAATTAATATCATCTTCTACTGAGTTTG 8533
QY 7501 GGAATCGCAGGTCCTTGGCTTCGCGATCAGAGACAGTGGCAGTTACAAATTAATAAGAAAT 7560
DB 8534 GGAATCGCAGGTCCTTGGCTTCGCGATCAGAGACAGTGGCAGTTACAAATTAATAAGAAAT 8593
QY 7561 TCOCGAGTTACTGTTGAAAAATCTTTTGTGAAGCAGCTGAAAGTAAATGTTAGATCG 7620
DB 8594 TCOCGAGTTACTGTTGAAAAATCTTTTGTGAAGCAGCTGAAAGTAAATGTTAGATCG 8653
QY 7621 GATATTACAGAAAAATGTTGCTTTAAACAGATATCAAGGCTCTGTAGAGACATTTGGAAATTA 7680
DB 8654 GATATTACAGAAAAATGTTGCTTTAAACAGATATCAAGGCTCTGTAGAGACATTTGGAAATTA 8713
QY 7681 GAGAGTGCCTATGACGAATTAATTTCTAATGGAAGATCAATATATCAGTATTAATAATTTCT 7740
DB 8714 GAGAGTGCCTATGACGAATTAATTTCTAATGGAAGATCAATATATCAGTATTAATAATTTCT 8773
QY 7741 AAGCTATTAGAAAAAATTTGATGTTATTTGTAAGAAATTAATCGGAATTTGAGAGCGAA 7800
DB 8774 AAGCTATTAGAAAAAATTTGATGTTATTTGTAAGAAATTAATCGGAATTTGAGAGCGAA 8833

QY 7801 GCAAAAGATTAACCGTAGAGACGGTAGCTGCGAGGCCATTAATCTCAAAAGCAAAAGAT 7860
DB 8834 GCAAAAGATTAACCGTAGAGACGGTAGCTGCGAGGCCATTAATCTCAAAAGCAAAAGAT 8893
QY 7861 GAAATGAATTCAGAGGTTGAAATTTGAGAAAGATATTTTCAATGAAAGAAATAGTAATCT 7920
DB 8894 GAAATGAATTCAGAGGTTGAAATTTGAGAAAGATATTTTCAATGAAAGAAATAGTAATCT 8953
QY 7921 AGCCCTCTTAAAGGAATTTGAGAAAGAAATTCATATGTCAAAGTGGAAAAAGAAACAGAGTG 7980
DB 8954 AGCCCTCTTAAAGGAATTTGAGAAAGAAATTCATATGTCAAAGTGGAAAAAGAAACAGAGTG 9013
QY 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGCAGGCGCAGAAATTAATTTCCGAA 8040
DB 9014 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGCAGGCGCAGAAATTAATTTCCGAA 9073
QY 8041 GCAAAAGATGCCGGAAGCTCTTAATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTTTT 8100
DB 9074 GCAAAAGATGCCGGAAGCTCTTAATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTTTT 9133
QY 8101 CATGAGATTAATGTAATATGGAAGCAACATTAATAAGAAATGAACAGAGTTCTTAA 8160
DB 9134 CATGAGATTAATGTAATATGGAAGCAACATTAATAAGAAATGAACAGAGTTCTTAA 9193
QY 8161 GCAGTAAACAGGTTCTGTATTTGGAGAGAGTTGAGTCAACCAAGGCAAGCTTACTGCA 8220
DB 9194 GCAGTAAACAGGTTCTGTATTTGGAGAGAGTTGAGTCAACCAAGGCAAGCTTACTGCA 9253
QY 8221 GGTAAACTATGTGTGAAGTTGAGAGAGAAATTTGTTCAAGAACAAATCGATTTGAATGCA 8280
DB 9254 GGTAAACTATGTGTGAAGTTGAGAGAGAAATTTGTTCAAGAACAAATCGATTTGAATGCA 9313
QY 8281 ATTTCTTAAGTGAAGGTTTGGATGAATGAATGAATGAATGCTTCTGTAATATCA 8340
DB 9314 ATTTCTTAAGTGAAGGTTTGGATGAATGAATGAATGAATGCTTCTGTAATATCA 9373
QY 8341 GAAAAATGAGAGAGAAATTTCCGAGACAGAGTGAATTAATCTTCAACAGCAAAAGTAATCT 8400
DB 9374 GAAAAATGAGAGAGAAATTTCCGAGAGCAGAGTGAATTAATCTTCAACAGCAAAAGTAATCT 9433
QY 8401 GAATCCGTAAGTTCTGTTACGAAAGCAGATTAATGAATAATGATTACAAAAAAATAT 8460
DB 9434 GAATCCGTAAGTTCTGTTACGAAAGCAGATTAATGAATAATGATTACAAAAAAATAT 9493
QY 8461 ATTTGAGAGTCATAGCTCTTGGCTTTAAATGATCAAGAAATGAAGCGAATTAAGATCT 8520
DB 9494 ATTTGAGAGTCATAGCTCTTGGCTTTAAATGATCAAGAAATGAAGCGAATTAAGATCT 9553
QY 8521 TTAGCGTAGCCGCTGTGCATGCAACAAAGAAACAAACAAGCATTTACAGATCAAAACAG 8580
DB 9554 TTAGCGTAGCCGCTGTGCATGCAACAAAGAAACAAACAAGCATTTACAGATCAAAACAG 9613
QY 8581 TTAACCTTACACCTGTATAATGAGAGAAACGTAATTCAACTTCGTGCAAAAAGCTTTGGCT 8640
DB 9614 TTAACCTTACACCTGTATAATGAGAGAAACGTAATTCAACTTCGTGCAAAAAGCTTTGGCT 9673
QY 8641 AAAAAATGAATTTATGAAATGTAAAGAACTGAGAGAGCTTGTGCGAGCGGAAACA 8700
DB 9674 AAAAAATGAATTTATGAAATGTAAAGAACTGAGAGAGCTTGTGCGAGCGGAAACA 9733
QY 8701 GCAGCCGTTGAAATTTATACAAAGAGTACTCAGAGAGATTTGTTGAGAGAAATTTGGAA 8760
DB 9734 GCAGCCGTTGAAATTTATACAAAGAGTACTCAGAGAGATTTGTTGAGAGAAATTTGGAA 9793
QY 8761 ATTGAGATTAATTAAGAAACGATTTGCAAGATTAATCGATTTGAAAGTCAACGAGAC 8820
DB 9794 ATTGAGATTAATTAAGAAACGATTTGCAAGATTAATCGATTTGAAAGTCAACGAGAC 9853
QY 8821 GGAACCAAGAGAGCTTGTGCGAAAGAAATGGAATTTCTGTGAAATTAATCAATTTCAAGG 8880
DB 9854 GGAACCAAGAGAGCTTGTGCGAAAGAAATGGAATTTCTGTGAAATTAATCAATTTCAAGG 9913
QY 8881 GAAACCAAAATCATCATTTGAAGATTAAGCAGAAATTTGGAACCGGAAGTGAATGTA 8940

```

Db 9914 GAAACAAATATCATTCATGAGATTAAGCCAGAAATTTGGAACCGGAAAGTGTAAATGTA 9973
Qy 8941 GATGCTTTGAATGAATCTGATGTAGATCTACAGGAAAAAGTGTGCTATGTGTGAAT 9000
Db 9974 GATGCTTTGAATGAATCTGATGTAGATCTACAGGAAAAAGTGTGCTATGTGTGAAT 10033
Qy 9001 GGTATTTGAAATGTTGATGTAAATATGATGATTAAGAAAAATGTGAAGCCAAAATCGGA 9060
Db 10034 GGTATTTGAAATGTTGATGTAAATATGATGATTAAGAAAAATGTGAAGCCAAAATCGGA 10093
Qy 9061 AGACATGCTATTTGTAAGAACTACCTGAGAAACAAAGATATCAAGATTTACAGAGCAAAA 9120
Db 10094 AGACATGCTATTTGTAAGAACTACCTGAGAAACAAAGATATCAAGATTTACAGAGCAAAA 10153
Qy 9121 GTAAATATTTCTTGGAAAAAGAGACGCTGAGCTGAGCTGCAATATCGAATGTACACAT 9180
Db 10154 GTAAATATTTCTTGGAAAAAGAGACGCTGAGCTGAGCTGCAATATCGAATGTACACAT 10213
Qy 9181 TCCAAATGAGATGATTTAAATAATTTGGCAAAAGCATGATGATCTTTCTCAATTAATACC 9240
Db 10214 TCCAAATGAGATGATTTAAATAATTTGGCAAAAGCATGATGATCTTTCTCAATTAATACC 10273
Qy 9241 AAAAATTCAAAAAATTAATTTTACTTTAGCATCAATGATGATGCAATGTGAATGTGAT 9300
Db 10274 AAAAATTCAAAAAATTAATTTTACTTTAGCATCAATGATGATGCAATGTGAATGTGAT 10333
Qy 9301 GGGGTGCTGAAGCAAGAGTGCAGAGCCAAAGCAAGTATGATTAAGATCAAAATA 9360
Db 10334 GGGGTGCTGAAGCAAGAGTGCAGAGCCAAAGCAAGTATGATTAAGATCAAAATA 10393
Qy 9361 AATAGAACTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAAAATCAATGTA 9420
Db 10394 AATAGAACTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAAAATCAATGTA 10453
Qy 9421 TATGCCGATATGATTAATAATTTAATTAATTAAGTAAAGCAAAATTTCTAAGGCTTTGGGAT 9480
Db 10454 TATGCCGATATGATTAATAATTTAATTAATTAAGTAAAGCAAAATTTCTAAGGCTTTGGGAT 10513
Qy 9481 GCCAAAGTCATGTCAGCTGCTTCCGCAACTGCCACTATTTGAAAAAATGAAATGTA 9540
Db 10514 GCCAAAGTCATGTCAGCTGCTTCCGCAACTGCCACTATTTGAAAAAATGAAATGTA 10573
Qy 9541 TTTAATTAATGCGATCGAGAAATTTAAATAATTTCTGCAAGATTTGAGAGGAAAGCTAAT 9600
Db 10574 TTTAATTAATGCGATCGAGAAATTTAAATAATTTCTGCAAGATTTGAGAGGAAAGCTAAT 10633
Qy 9601 AAAAAAAGCTCGGTAGAGATCTAATCAGTGAAGTGTATACGATTAATATACATGGCAT 9660
Db 10634 AAAAAAAGCTCGGTAGAGATCTAATCAGTGAAGTGTATACGATTAATATACATGGCAT 10693
Qy 9661 TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATCAAGAGAGAAAAAGGAAA 9720
Db 10694 TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATCAAGAGAGAAAAAGGAAA 10753
Qy 9721 AAATGA 9726
Db 10754 AAATGA 10759

```

```

RESULT 3
AA167642
ID AA167642 standard; DNA; 2780 BP.
XX
AC AA167642;
XX
DT 27-FEB-2002 (first entry)
XX
DE F. necrophorum truncated leukotoxin protein (SX) encoding DNA.
XX
KM Leukotoxin; infection; immunisation; liver abscess; foot rot;
XX bactericide; vaccine; truncated; SX; db.
XX

```

```

OS Fusobacterium necrophorum.
FH Key Location/Qualifiers
FT CDS 1..2780
FT /tag=a
FT /product="truncated leukotoxin (SX)"
XX
PN W0200180886-A2.
XX
PD 01-NOV-2001.
XX
PE 25-APR-2001; 2001WO-US013240.
XX
PR 25-APR-2000; 2000US-00558257.
XX 24-APR-2001; 2001US-00841786.
XX
PA (UNIV ) UNIV KANSAS STATE RES FOUND.
XX
PI Nagaraja TG, Stewart GC, Narayanan SK, Chengappa MW;
XX WPI: 2002-049245/06.
XX P-PSDB; AAG66007.
XX
DR Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an
DR animal against an infection e.g. foot rot, or liver abscesses caused by
XX the bacterium.
XX
PS Claim 10; Page 96-97; 108bp; English.
XX
CC The invention provides an isolated Fusobacterium necrophorum leukotoxin
CC polypeptide and its truncated versions. The leukotoxin protein can be
CC expressed by standard recombinant methodology. The leukotoxin gene is
CC useful for preparing a vaccine which confers effective immunity against
CC infection caused by F. necrophorum. The vaccine is useful for immunising
CC an animal against liver abscesses caused by F. necrophorum and for
CC preventing foot rot caused by F. necrophorum infection. The present
CC sequence represents a DNA encoding a F. necrophorum truncated leukotoxin
CC protein (SX)
XX
SQ Sequence 2780 BP; 1031 A; 362 C; 637 G; 750 T; 0 U; 0 Other;
XX
Qy 919 GGGAGATTATTAAGAAAGAGGAGGAAAAAGAACTTATACATCTCTTAAGTTATCA 978
Db 1 GGGAGATTATTAAGAAAGAGGAGGAAAAAGAACTTATACATCTCTTAAGTTATCA 60
Qy 979 GATGTGAAGCTTCGTAAGAGTAAATTAAGAAAAAGTATGAAAGAAATGTTGACATT 1038
Db 61 GATGTGAAGCTTCGTAAGAGTAAATTAAGAAAAAGTATGAAAGAAATGTTGACATT 120
Qy 1039 AAGCTGAAGCAAGAAATTTCTATGATGCACTTATGTTACTAGCTTCAAGACACTCT 1098
Db 121 AAGCTGAAGCAAGAAATTTCTATGATGCACTTATGTTACTAGCTTCAAGACACTCT 180
Qy 1099 TTTAGCTTTGTACAGTCTTATTTCTCTCTCAATTTAAATGATTTTAAAGTTATG 1158
Db 181 TTTAGCTTTGTACAGTCTTATTTCTCTCTCAATTTAAATGATTTTAAAGTTATG 240
Qy 1159 ACAAGTAAGTCCAGTGTCTTATTGAAAAAGATGCCAAAGTGAAGCAACAGAAAGAAAG 1218
Db 241 ACAAGTAAGTCCAGTGTCTTATTGAAAAAGATGCCAAAGTGAAGCAACAGAAAGAAAG 300
Qy 1219 GCAATATTCATTTCTTACAGTGAAGTGAAGCAACTATGAGAGAGAGTCTTCCATTA 1278
Db 301 GCAATATTCATTTCTTACAGTGAAGTGAAGCAACTATGAGAGAGAGTCTTCCATTA 360
Qy 1279 AAAATTAACCAATTTATTTATTTGAGAAAGCAATGAAACCTTCTCAGTATCGAGCGGGA 1338
Db 361 AAAATTAACCAATTTATTTATTTGAGAAAGCAATGAAACCTTCTCAGTATCGAGCGGGA 420

```

OY	1339	TATATTTCTGCAAAAAGTAATTTCCAACTGTAATCTATTGAAGGAAAGTAAATCCGAGGA	1338
Db	421	TATATTTCTGCAAAAAGTAATTTCCAACTGTAATCTATTGAAGGAAAGTAAATCCGAGGA	480
OY	1399	AGAGCAGATATTACTTCTCAAAATCTGAAATATCTATTGATGCTTCGTTTCGTTGGAAAG	1458
Db	481	AGAGCAGATATTACTTCTCAAAATCTGAAATATCTATTGATGCTTCGTTTCGTTGGAAAG	540
OY	1459	ATGAGAGATTCCAATPAAAGTAGCTCTTTCAGTATTTGTGACGGAAGGAAATPAAATCT	1518
Db	541	ATGAGAGATTCCAATPAAAGTAGCTCTTTCAGTATTTGTGACGGAAGGAAATPAAATCT	600
OY	1519	TCCGTCAGATTTGCTPAAAGACAAAGTAGAATACGAACGGATGATGTAAATGTGAGA	1578
Db	601	TCCGTCAGATTTGCTPAAAGACCAAAAGTAGAATACGAACCGAATGATGTAAATGTGAGA	660
OY	1579	AGTGAAGGATTTAATTCGATTCGAGCTGCTGAAAAGGTGGAATTGGGGGATAGTGTAT	1638
Db	661	AGTGAAGGATTTAATTCGATTCGAGCTCTGAAAAGGTGGAATTGGGGGATAGTGTAT	720
OY	1639	GGGGTTGTGCTGCAAAATATTTCTAACTATAATGCTTCTCCCGTATAGATGTAGTGA	1698
Db	721	GGGGTTGTGCTGCAAAATATTTCTAACTATAATGCTTCTCCCGTATAGATGTAGTGA	780
OY	1699	TATCTACATGCCAAGACGCACTAAATGTGAGGCTCATPAACTACTPAAAATATGTCT	1758
Db	781	TATCTACATGCCAACAAGACGCACTAAATGTGAGGCTCATPAACTACTPAAAATATGTCT	840
OY	1759	CTGCAAAACAGGATCTGATTTGGGAATCTCAAGTTTATGANTGATCAGGTTTATGAATCA	1818
Db	841	CTGCAAAACAGATCTGATTTGGGAATCTTCAAGTTTATGAAATGATCAGGTTTATGAATCA	900
OY	1819	GATCATCTPAAATCAATTTTAGATGCAATPAAAACAGCGTTTGGAGAGACAGTGTCAAT	1878
Db	901	GATCATCTPAAATCAATTTTAGATGCAATPAAAACAGCGGTTTGGAGAGACAGTGTCAAT	960
OY	1879	GAGCAAAATPAAAGATPAAAGCTAAACGAATTTATGTGTGGTGTCTGCAACCATAGCA	1938
Db	961	GAGCAAAATPAAAGAAATPAAAGCTAAACGAATTTATGTGTGGTGTCTGCAACCATAGCA	1020
OY	1939	AATCATTAATTTCTGCTTCTGTGGCAATAGAGAGATGTGAAAGCTTTCTTCAGAGTG	1998
Db	1021	AATCATTAATTTCTGCTTCTGTGGCAATAGAGAGATGTGAAAGCTTTCTTCAGAGATG	1080
OY	1999	GAAAGGATTAATGTAAAGGCAATTAATGAAGCTCAAAATCTTCGAGCACTACGTCAGT	2058
Db	1081	GAAAGGATTAATGTAAAGGCAATTAATGAAGCTCAAAATCTTCGAGCACTACGTCAGT	1140
OY	2059	GGAAGTGTGCTGTACGAAGAGAGAAAAAAGAACTTATTTGGAATGACAGAGTTTT	2118
Db	1141	GGAAGTGTGCTGTACGAAGAGAGAAAAAAGAACTTATTTGGAATGACAGAGTTTT	1200
OY	2119	TATGGAATCTATPAAAAATPAAATGCTTCTGTGACAATGCCGATCATGCTGAATTTGATCG	2178
Db	1201	TATGGAATCTATPAAAAATPAAATGCTTCTGTGACAATTTGCGATTCAGTCGAATTTGATCG	1260
OY	2179	GAAAGAAATTAATGATATCAACAGTGAATPAAAAATTTGAATATPAAAAATCTTCGAAAAATG	2238
Db	1261	GAAAGAAATTAATGATATCAACAGTGAATPAAAAATTTGAATATPAAAAATCTTCGAAAAATG	1320
OY	2239	GCAAAGTCTGTATTGATPAAATTTAGACTTTTAAAGAGAGCTTTTGGAAAAAGAAAGAA	2298
Db	1321	GCAAAGTCTGTATTGATPAAATTTAGACTTTTAAAGAGAGCTTTTGGAAAAAGAAAGAA	1380
OY	2299	ACTCCAGAAATATGATCCGAAAGATATTGAATCTATTGAAAAATTAATTGAATGCAATTTCA	2358
Db	1381	ACTCCAGAAATATGATCCGAAAGATATTGAATCTATTGAAAAATTAATTGAATGCAATTTCA	1440
OY	2359	GAAAAAATTTGATGAGAAAAACCGAGCTTTTATCTPAAATTTGGTGAAGATACATTAATCTCT	2418
Db	1441	GAAAAAATTTGATGAGAAAAACCGAGCTTTTATCTPAAATTTGGTGAAGATACATTAATCTCT	1500
OY	2419	CCGAGTGAACCTTCAAAAAACAGAACTGCTATAGAAATTTGCAACATATGTTCAAGGAGAA	2478

Db	1501	CCGGATGGAACCTTCAAAAAACGGAACGTCTATAGAAATTGCAAACTATGTTCAAGAGAA	1566
Oy	2479	ATGAAAAAATTAGAGAAAAATTACCGAAAAGATTTAAAGCTTTTTCAGAAAGATTGAGT	2538
Db	1561	ATGAAAAAATTAGAGAAAAATTACCGAAAGATTTTAAAGCTTTTTCAGAAAGATTGAGT	1620
Oy	2539	GGACTGATTAAGAAACCTTGAAATTTTACAGAGTAGAAATTATGCAATTTTTCACACT	2598
Db	1621	GGACTGATTAAGAAACCTTGAAATTTTACAGAGTAGAAATTATGCAATTTTTCACACT	1680
Oy	2599	TTTACCTTCCTCCGAGCTAATGGAAGAAAGATGTTCTTCTGAGGAGAGCTGTTTCG	2658
Db	1681	TTTACCTTCCTCCGAGCTAATGGAAGAAAGATGTTCTTCTGAGGAGAGCTGTTTCG	1740
Oy	2659	TGGGTAGAACAGAGAAATTAAGCAAGGTATCCGTTGGAAAAAGAGCTTAACTTGCTGCA	2718
Db	1741	TGGGTAGAACAGAGAAATTAAGCAAGGTATCCGTTGGAAAAAGAGCTTAACTTGCTGCA	1800
Oy	2719	AAAAAAGATTTTAAATATAAAGCTATCAATAAGCAGAAACAGTAAATTTAGTTGAAAT	2778
Db	1801	AAAAAAGATTTTAAATATAAAGCTATCAATAAGCAGAAACAGTAAATTTAGTTGAAAT	1860
Oy	2779	ATTGGACTTGGAGAGACAGTACATCCGGAAGTGCAGTCGGAAGAAATTTAAATGTTCAA	2838
Db	1861	ATTGGACTTGGAGAGACAGTACATCCGGAAGTGCAGTCGGAAGAAATTTAAATGTTCAA	1920
Oy	2839	AGATCGAAAAATTCAGGCTATCGTAGAAGCTTAAAGAAAAAGCTGAAATTAACAGAGAAAAAT	2898
Db	1921	AGATCGAAAAATTCAGGCTATCGTAGAAGCTTAAAGAAAAAGCTGAAATTAACAGAGAAAAAT	1980
Oy	2899	ATTATATCAGATGCATTTGAACAGACACTTTTTCATGTAGCCGAGATCTTTTAAATGTTGCTCA	2958
Db	1981	ATTATATCAGATGCATTTGAACAGACACTTTTTCATGTAGCCGAGATCTTTTAAATGTTGCTCA	2040
Oy	2959	GGTGGAAATGCATCAATGGAATGGGAAGTTATAGTAGAGTATCAGTAAGGCAAGACTT	3018
Db	2041	GGTGGAAATGCATCAATGGAATGGGAAGTTATAGTAGAGTATCAGTAAGGCAAGACTT	2100
Oy	3019	TCCATTATATGCGAAGCATATTGAAAAAGCTTAAATAAAATTTGCTTTAAACAGTAAGAT	3078
Db	2101	TCCATTATATGCGAAGCATATTATTAAAGCTTATAAAAAATTTGCTTTAAACAGTAAGAT	2160
Oy	3079	GATACCTCTGTTTGGAATGCTGCCGGTTCAGCCGGAAATCGGAACGAAAAATGCCGCGCTC	3138
Db	2161	GATACCTCTGTTTGGAATGCTGCCGGTTCAGCCGGAAATCGGAACGAAAAATGCCGCGCTC	2220
Oy	3139	GGGGTTGCTGTGCGGTAAATGATTATGATATTTCAACAAAGCTTCCATTGAAGATAT	3198
Db	2221	GGGGTTGCTGTGCGGTAAATGATTATGATATTTCAACAAAGCTTCCATTGAAGATAT	2280
Oy	3199	GACGAAGGACAAATTAATATGATTAAGATTAADATGTGAAGTACAGTAATGCCGGA	3258
Db	2281	GACGAAGGACAAATTAATATGATTAAGATTAADATGTGAAGTACAGTAATGCCGGA	2340
Oy	3259	TCTTTAGAAGTAGTGCAAAAACGACCGGAACAATCAACAGTATTTCTGTTCCCGAGGA	3318
Db	2341	TCTTTAGAAGTAGTGCAAAAACGACCGGAACAATCAACAGTATTTCTGTTCCCGAGGA	2400
Oy	3319	ATTAAATAAGTTGGAAGTAAACGAGTGAAGAAAAACCGAAATCAAGAAAGAACACAGAG	3378
Db	2401	ATTAAATAAGTTGGAAGTAAACGAGTGAAGAAAAACCGAAATCAAGAAAGAACACAGAG	2460
Oy	3379	GGATTTTTTGGCAAAATCGGAAACAAAGTGAATCTGTGTAATAAATTTAAATTTACCGATAGT	3438
Db	2461	GGATTTTTTGGCAAAATCGGAAACAAAGTGAATCTGTGTAATAAATTTAAATTTACCGATAGT	2520
Oy	3439	ATGGAATTCATTAACAGAAAAATTTCAAAATTAACATTTCTGAAGAGTAAAAAAAGCGGAG	3498
Db	2521	ATGGAATTCATTAACAGAAAAATTTCAAAATTAACATTTCTGAAGAGTAAAAAAAGCGGAG	2580
Oy	3499	AATCTTCTCTGGAACGTTTCTCATACTCCGATTAAGAGACCGTCTTTCAGTTTGGAGCT	3558

```
Db      2581 AATCTTCCTTGAAGCTTCTCATACTCCGATAAAGACCGTCTTTTCACTTTGGAGCT 2640
Oy      3559 TCTGGAAGTCTTTCTTCAATAATATTTAAAAAGAAACATCTGCTGTCTAGATGAGTA 3618
Db      2641 TCTGGAAGTCTTTCTTCAATAATATTTAAAAAGAAACATCTGCTGTCTAGATGAGTA 2700
Oy      3619 AAGATAAATTTTGAAGGACCAATTAAGATGAGAGTACTTCTTCTGATCTACTTT 3678
Db      2701 AAGATAAATTTTGAAGGACCAATTAAGATGAGAGTACTTCTTCTGATCTACTTT 2760
Oy      3679 GTTGAGCATGGGCGGATC 3698
Db      2761 GTTGAGCATGGGCGGATC 2780

RESULT 4
AA167645
ID      AA167645 standard; DNA; 2322 BP.
XX
AC      AA167645;
XX
DT      27-FEB-2002 (first entry)
XX
DE      F. necrophorum truncated leukotoxin protein (FINAL) encoding DNA.
XX
KW      Leukotoxin; infection; immunisation; liver abscess; foot rot;
XX      bactericide; vaccine; truncated; FINAL; db.
XX
OS      Fusobacterium necrophorum.
XX
FH      Key Location/Qualifiers
FT      CDS 1..2322
FT      /tag= a
FT      /product= "truncated leukotoxin (FINAL)"
XX
PN      MO200180886-A2.
XX
PD      01-NOV-2001.
XX
PF      25-APR-2001; 2001WO-US013240.
XX
PR      25-APR-2000; 2000US-00558257.
XX      24-APR-2001; 2001US-00841786.
XX
PA      (UNIV ) UNIV KANSAS STATE RES FOUND.
XX
PI      Nagendra TG, Stewart GC, Narayanan SK, Chengappa MM;
XX
DR      WPI, 2002-049245/06.
XX
DR      P-PSDB; AAG66010.
XX
PT      Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an
XX      animal against an infection e.g. foot rot, or liver abscesses caused by
XX      the bacterium.
XX
PS      Claim 10; Page 100-101; 108bp; English.
XX
CC      The invention provides an isolated Fusobacterium necrophorum leukotoxin
XX      polypeptide and its truncated versions. The leukotoxin protein can be
XX      expressed by standard recombinant methodology. The leukotoxin gene is
XX      useful for preparing a vaccine which confers effective immunity against
XX      infection caused by F. necrophorum. The vaccine is useful for immunising
XX      an animal against liver abscesses caused by F. necrophorum and for
XX      preventing foot rot caused by F. necrophorum infection. The present
XX      sequence represents a DNA encoding a F. necrophorum truncated leukotoxin
XX      protein (FINAL)
XX
SQ      Sequence 2322 BP; 921 A; 299 C; 532 G; 570 T; 0 U; 0 Other;
Query Match 23.9%; Score 2322; DB 6; Length 2322;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      7405 GGAATTAAGCCATGTGAAGACTTCTGATATTAATGGAAAAAATGATGATATTAACA 7464
Db      1 GGAATTAAGCCATGTGAAGACTTCTGATATTAATGGAAAAAATGATGATATTAACA 60
Oy      7465 GAGGACAAAGATTAATATACCTTCTACTGTGTGTTGGAACTGACGCTTGTCCGCA 7524
Db      61 GAGGACAAAGATTAATATACCTTCTACTGTGTGTTGGAACTGACGCTTGTCCGCA 120
Oy      7525 TCAGAACAGTGTGCGAGTTACAAATTAATAAAGAAATTCGAGATTACGTTGAAATCT 7584
Db      121 TCAGAACAGTGTGCGAGTTACAAATTAATAAAGAAATTCGAGATTACGTTGAAATCT 180
Oy      7585 TTTGTAAAGACAGCTGAATAAGTAAATGTTAGATCGATATTAACAGAAATGTTCTTTA 7644
Db      181 TTTGTAAAGACAGCTGAATAAGTAAATGTTAGATCGATATTAACAGAAATGTTCTTTA 240
Oy      7645 ACAGCATATCAAGGCTCTGTAGAGCATTTGGAAATAGAGCTGCTATGCAAAATTAAT 7704
Db      241 ACAGCATATCAAGGCTCTGTAGAGCATTTGGAAATAGAGCTGCTATGCAAAATTAAT 300
Oy      7705 TCTAATGCAAGATCAAAATATTCAGTATTAATAATTTCTAAGCTATTGAGAAAAATATGAT 7764
Db      301 TCTAATGCAAGATCAAAATATTCAGTATTAATAATTTCTAAGCTATTGAGAAAAATATGAT 360
Oy      7765 GTTATTTGTAAGATTAATAATCGGAATTGAGAGCGGAAGCAAAAGATTAACCGTAGAGCG 7824
Db      361 GTTATTTGTAAGATTAATAATCGGAATTGAGAGCGGAAGCAAAAGATTAACCGTAGAGCG 420
Oy      7825 GTAGTGTCCGAGCCATTATCTCAAAAGCAAAATGAAATGAAATTCAGAGTTGAAAT 7884
Db      421 GTAGTGTCCGAGCCATTATCTCAAAAGCAAAATGAAATGAAATTCAGAGTTGAAAT 480
Oy      7885 GAGAAAGATATTTTCAATGAAGAAAAATAGAGTAATCAGCCCTTCTAAAGAAATTTGAAAG 7944
Db      481 GAGAAAGATATTTTCAATGAAGAAAAATAGAGTAATCAGCCCTTCTAAAGAAATTTGAAAG 540
Oy      7945 GAAATCAATGTCAAAAGTGAAGAAAAAGCAAGTGAATCTGAATCTCAAGAGCTTCT 8004
Db      541 GAAATCAATGTCAAAAGTGAAGAAAAAGCAAGTGAATCTGAATCTCAAGAGCTTCT 600
Oy      8005 GTAGAGCAGTAGCAGAGGCGCAGAAATTAATTTCCGAAGCAAAAGATGCCGAAGCTTAT 8064
Db      601 GTAGAGCAGTAGCAGAGGCGCAGAAATTAATTTCCGAAGCAAAAGATGCCGAAGCTTAT 660
Oy      8065 TTGAAAGTTAGTACAAATCCGAAAGAGTATTTTTCATGCAATTAATGTGAATATGAA 8124
Db      661 TTGAAAGTTAGTACAAATCCGAAAGAGTATTTTTCATGCAATTAATGTGAATATGAA 720
Oy      8125 GCAACACATTAATAAGTAAACAGAGTTCTAAAGCAGTAACAGGTTCTGATTTGGGA 8184
Db      721 GCAACACATTAATAAGTAAACAGAGTTCTAAAGCAGTAACAGGTTCTGATTTGGGA 780
Oy      8185 GGAGTTGAGTACCAAGGACAGAGTACTGTGCGAGTAAACTATGTAGAGTTGAG 8244
Db      781 GGAGTTGAGTACCAAGGACAGAGTACTGTGCGAGTAAACTATGTAGAGTTGAG 840
Oy      8245 GAAGCAATTTGTTCAAGAACAAATCGATTTGAAATGCAATTTCTAAGATTAAGAGTTGGAT 8304
Db      841 GAAGCAATTTGTTGCAAGAACAAATCGATTTGAAATGCAATTTCTAAGATTAAGAGTTGGAT 900
Oy      8305 GAAGATTAAGTAACTAGCTTAATCTTCTGTAGTACGAAATGAGAGAGAAATTCGCGGA 8364
Db      901 GAAGATTAAGTAACTAGCTTAATCTTCTGTAGTACGAAATGAGAGAGAAATTCGCGGA 960
Oy      8365 GCAGAGTAAATACCTTCTACAGACAAAGTAAATCTGATTCGTTTACGAAG 8424
Db      961 GCAGAGTAAATACCTTCTACAGACAAAGTAAATCTGATTCGTTTACGAAG 1020
Oy      8425 CAAAGATTATGAATAATATGATTAACAAATAAATATTTTCAGAGTCAATGCTCTTGT 8484
Db      1021 CAAAGATTATGAATAATATGATTAACAAATAAATATTTTCAGAGTCAATGCTCTTGT 1080
Oy      8485 TTAATGATTAACAAGAAATGAAGCAATATGAAATCTTTAGCGGTAGCCGATGTGATGCA 8544
```

```
Db 1081 TTTAAATGATACAAAGAAATGAAAGCAATATAGAAATCTTTAGCGGTAGCCGGTGTGATGCA 1140
Qy 8545 CAAGGAAACAAAGAAAGCAATTTACGAGATCAACCAAGTTTACTTTCACATCTTAATGGA 8604
Db 1141 CAAGGAAACAAAGCAATTTACGAGATCAACCAAGTTTACTTTCACATCTTAATGGA 1200
Qy 8605 GGAAGACGTATCTCAACTTGTGTCAGAAAGCTTGGCTAAATGAAATGAAATGAAATGGA 8664
Db 1201 GGAAGACGTATCTCAACTTGTGTCAGAAAGCTTGGCTAAATGAAATGAAATGGA 1260
Qy 8665 AAAGGAACTGGAGAGACCTTACTCGAGCGGAAACAGACGCGCTTGAATTTATCAAG 8724
Db 1261 AAAGGAACTGGAGAGACCTTACTCGAGCGGAAACAGACGCGCTTGAATTTATCAAG 1320
Qy 8725 ACTATCAAGAGAGCAATTTGGTTCAGAGAAATTTGGGAAATTTGGAGATTAATTAAGAAAT 8784
Db 1321 ACTATCAAGAGAGCAATTTGGTTCAGAGAAATTTGGGAAATTTGGAGATTAATTAAGAAAT 1380
Qy 8785 GCAAGAGATTAATCAAGATTTGTAAGATCAACGAGAGCGGAAACCAAGAGAGCTTGTGCGA 8844
Db 1381 GCAAGAGATTAATCAAGATTTGTAAGATCAACGAGAGCGGAAACCAAGAGAGCTTGTGCGA 1440
Qy 8845 AAGATGATTTCTGTGTAATTAATTTCAAGGAGAAACAAATTCATCCATTTGAAGAT 8904
Db 1441 AAGATGATTTCTGTGTAATTAATTTCAAGGAGAAACAAATTCATCCATTTGAAGAT 1500
Qy 8905 AAAGCCAGAAATTTGGGAAACCGGAGGTGTAATTTGTAAGATTTGTAAGATTTGTAAGAT 8964
Db 1501 AAAGCCAGAAATTTGGGAAACCGGAGGTGTAATTTGTAAGATTTGTAAGATTTGTAAGAT 1560
Qy 8965 GATCTACAAAGAGAAAGATGCTGCTATGCTGGAATTTGTAAGATTTGTAAGATTTGTAAGAT 9024
Db 1561 GATCTACAAAGAGAAAGATGCTGCTATGCTGGAATTTGTAAGATTTGTAAGATTTGTAAGAT 1620
Qy 9025 AATGTGATTAAGAAATTTGTAAGAGCAAAATCGAAGACATGCTATTTGTAAGATTTGTAAGAT 9084
Db 1621 AATGTGATTAAGAAATTTGTAAGAGCAAAATCGAAGACATGCTATTTGTAAGATTTGTAAGAT 1680
Qy 9085 GGAAGAAACAAAGATTTCAAGCAATTTCAAGAGCAAAAGTAATTTCTGGAAGAAAGAGAC 9144
Db 1681 GGAAGAAACAAAGATTTCAAGCAATTTCAAGAGCAAAAGTAATTTCTGGAAGAAAGAGAC 1740
Qy 9145 GCTGACGTGACGTGCAATATCGAATGTAACATTTCCATGAGATGGAATTTAAAT 9204
Db 1741 GCTGACGTGACGTGCAATATCGAATGTAACATTTCCATGAGATGGAATTTAAAT 1800
Qy 9205 TTGGCAAGACGATGATCTTCTCAATTAATTAACCAAAATTTCAAAATTAATTTACT 9264
Db 1801 TTGGCAAGACGATGATCTTCTCAATTAATTAACCAAAATTTCAAAATTAATTTACT 1860
Qy 9265 TTAGCATCAAGTAGAATCGAATGTAATGTTCAATGGGGTGGCTGAAGCAAGAGTGA 9324
Db 1861 TTAGCATCAAGTAGAATCGAATGTAATGTTCAATGGGGTGGCTGAAGCAAGAGTGA 1920
Qy 9325 GAGACCAAGCAAGTAGAATCGAATGTAATGTAATTAATTAAGATTAATTAATTTGATTTA 9384
Db 1921 GAGACCAAGCAAGTAGAATCGAATGTAATGTAATTAATTAATTAATTTGATTTA 1980
Qy 9385 GCAAGGAAATTTAAACAGAGGAAACATCAATGTAATGCGGATATGATTAATTAATTTAAT 9444
Db 1981 GCAAGGAAATTTAAACAGAGGAAACATCAATGTAATGCGGATATGATTAATTAATTTAAT 2040
Qy 9445 AATTAATTAAGAGCAAAATTTCTAAGGCTATTTGGGAGCCCAAAAGTCATGCTCAGCTGCT 9504
Db 2041 AATTAATTAAGAGCAAAATTTCTAAGGCTATTTGGGAGCCCAAAAGTCATGCTCAGCTGCT 2100
Qy 9505 TCGGCAACTGCACTATTTGAAAAATTAAGATTAATTTAATTAATGATCGGAGATTT 9564
Db 2101 TCGGCAACTGCACTATTTGAAAAATTAAGATTAATTTAATTAATGATCGGAGATTT 2160
Qy 9565 AAAAATTAATCTGGCAAGATTTGAAAGGAAAGCTAATTAATTAATTAATTAATTTAAT 9624
```

```
Db 2161 AAAATTAATCTGGCAAGATTTGAAAGGAAACCTAATTAATTAATTAATTAATTTAAT 2220
Qy 9625 CAGGTAGACTGTGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9684
Db 2221 CAGGTAGACTGTGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
Qy 9685 TTGACATATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9726
Db 2281 TTGACATATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2322

RESULT 5
AA167643
ID AA167643 standard; DNA; 2141 BP.
XX
AC AA167643;
XX
DT 27-FEB-2002 (first entry)
XX
DE F. necrophorum truncated leukotoxin protein (GAS) DNA.
XX
KW Leukotoxin; infection; immunisation; liver abscess; foot rot;
KW bactericide; vaccine; truncated; GAS; de.
XX
OS Fusobacterium necrophorum.
XX
PN MO200180886-A2.
XX
PD 01-NOV-2001.
XX
PF 25-APR-2001; 2001WO-US013240.
XX
PR 25-APR-2000; 2000US-00558257.
XX
PR 24-APR-2001; 2001US-008411786.
XX
PA (UNIT) UNIV KANSAS STATE RES FOUND.
XX
PI Nagendra TG, Stewart GC, Narayanan SK, Chengappa MW;
XX
DR MPI; 2002-049245/06.
XX
PT Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an
PT animal against an infection e.g. foot rot, or liver abscesses caused by
PT the bacterium.
XX
PS Claim 10; Page 97-98; 108pp; English.
XX
CC The invention provides an isolated Fusobacterium necrophorum leukotoxin
CC polypeptide and its truncated versions. The leukotoxin protein can be
CC expressed by standard recombinant methodology. The leukotoxin gene is
CC useful for preparing a vaccine which confers effective immunity against
CC infection caused by F. necrophorum. The vaccine is useful for immunising
CC an animal against liver abscesses caused by F. necrophorum and for
CC preventing foot rot caused by F. necrophorum infection. The present
CC sequence represents a DNA encoding a F. necrophorum truncated leukotoxin
CC protein (GAS)
XX
SQ Sequence 2141 BP; 763 A; 310 C; 513 G; 555 T; 0 U; 0 Other;

Query Match 22.0%; Score 2141; DB 6; Length 2141;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3553 GAGGCTTTGGAAGGTTCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3612
Db 1 GAGGCTTTGGAAGGTTCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 60
Qy 3613 GAGGAAAGATTAATTTGAAGAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3672
Db 61 GAGGAAAGATTAATTTGAAGAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
Qy 3673 ACTTTGTTGAGAGATGGGGCGGATCTGCTCACTTCAAGTGAATCATATTGGAAGTGA 3732
```

Db 121 ACTTTGTTGAGAGCATGGGGGCGGATCTGCTGCATTCAGTGGATCATATTGGAAGTGA 180
 Qy 3733 AATAGCAACATCAGTGTGTTTACCTGAGCGGCTGTAATTAATTAATTCAGAAATGA 3792
 Db 181 AATAGCAACATCAGTGTGTTTACCTGAGCGGCTGTAATTAATTAATTCAGAAATGA 240
 Qy 3793 ACAATGCTTTGGTTTAAAAATAGTGAATATGGAATGCCAATTAATTTAAAGTAATGCT 3852
 Db 241 ACMAATGCTTTGGTTTAAAAATAGTGAATATGGAATGCCAATTAATTTAAAGTAATGCT 300
 Qy 3853 TTGAGTGGAGGAACTCAAGTAGCAGCAGAGCAGGTTTGGAAAGCAGTTAAAGAAAGTGA 3912
 Db 301 TTGAGTGGAGGAACTCAAGTAGCAGCAGAGCAGGTTTGGAAAGCAGTTAAAGAAAGTGA 360
 Qy 3913 GGACAGAGGAAAAAGTTATCTATTGGGAACCTTCTGCTTATCACTTAAGTGAACATGA 3972
 Db 361 GGACAGAGGAAAAAGTTATCTATTGGGAACCTTCTGCTTATCACTTAAGTGAACATGA 420
 Qy 3973 GTTTCGCAAAATCAGAAAAATTAACATAGAGAGAGAAATCGAAAGCCAAAAAATGGAT 4032
 Db 421 GTTTCGCAAAATCAGAAAAATTAACATAGAGAGAGAAATCGAAAGCCAAAAAATGGAT 480
 Qy 4033 GTTGAATCAGCTGTTATCAAGCGGACACCAAGTGAACAGAGCTTTAAATTTACAAAGCT 4092
 Db 481 GTTGAATCAGCTGTTATCAAGCGGACACCAAGTGAACAGAGCTTTAAATTTACAAAGCT 540
 Qy 4093 GGAAAGTCAAAATGGAATCTGTAGGGGCTACTGTGACTGTGCCAAATTTAAACAACAAGTA 4152
 Db 541 GGAAAGTCAAAATGGAATCTGTAGGGGCTACTGTGACTGTGCCAAATTTAAACAACAAGTA 600
 Qy 4153 AATGCTTCTATTAGTGGTGGAGATTAATCAACGTTAATCGACCGGACCAAAAGCTCTT 4212
 Db 601 AATGCTTCTATTAGTGGTGGAGATTAATCAACGTTAATCGACCGGACCAAAAGCTCTT 660
 Qy 4213 TTAGCAACCACTCAAGTGACTGCTGCAATGAGAGAGGAGGCAATTAAGTTCTGGAGG 4272
 Db 661 TTAGCAACCACTCAAGTGACTGCTGCAATGAGAGAGGAGGCAATTAAGTTCTGGAGG 720
 Qy 4273 GGATTAGGAAATTAATCAAGGGGCTGTTTCTGTCAATTAAGATTGACAAATGACGTGAAGCT 4332
 Db 721 GGATTAGGAAATTAATCAAGGGGCTGTTTCTGTCAATTAAGATTGACAAATGACGTGAAGCT 780
 Qy 4333 AGCGTTGATTAATCTTCATCGAAGAGCTAATGAATCAATGTCAATTCGCAAAAGTGC 4392
 Db 781 AGCGTTGATTAATCTTCATCGAAGAGCTAATGAATCAATGTCAATTCGCAAAAGTGC 840
 Qy 4393 AAAAGAGTTCGATCTAGCAAAAGAAATACAGGCTTACTTAATGGAAGATTAAGAA 4452
 Db 841 AAAAGAGTTCGATCTAGCAAAAGAAATACAGGCTTACTTAATGGAAGATTAAGAA 900
 Qy 4453 TATTTAGAGATCGTGTATTAATAAGACTGGAATGTTATTAATCGAAGGAAACAATGA 4512
 Db 901 TATTTAGAGATCGTGTATTAATAAGACTGGAATGTTATTAATCGAAGGAAACAATGA 960
 Qy 4513 GAAAAAGCAAGAAAAAGAGAGCGTTCATTGTAATGCTGCTTTATCGTGTCTGA 4572
 Db 961 GAAAAAGCAAGAAAAAGAGAGCGTTCATTGTAATGCTGCTTTATCGTGTCTGA 1020
 Qy 4573 ACCGATTAATCCGCTGAGAGAGATTAATGCAATGCAATGCTTAAATTAATTAATTA 4632
 Db 1021 ACCGATTAATCCGCTGAGAGATTAATGCAATGCAATGCTTAAATTAATTAATTA 1080
 Qy 4633 GCAGATTGAGTGAAGCAATAAGAGCCGAGAGATTAATTAATGATCGAAGCAATGA 4692
 Db 1081 GCAGATTGAGTGAAGCAATAAGAGCCGAGAGATTAATTAATGATCGAAGCAATGA 1140
 Qy 4693 AATGTGAGGCAAAATCATCTACTGTTGTTGTAATGCGGCTTCTGACCTTATCAGC 4752
 Db 1141 AATGTGAGGCAAAATCATCTACTGTTGTTGTAATGCGGCTTCTGACCTTATCAGC 1200
 Qy 4753 AAAGATGCTTTTCAAGAAATGGAGTCTGAGCAATGAGCAAGTATTAACAATGACAGATT 4812
 Db 1201 AAAGATGCTTTTCAAGAAATGGAGTCTGAGCAATGAGCAAGTATTAACAATGACAGATT 1260

Qy 4813 GCAAGGATTAAGGAAATTTCTGCTGATTCCTTAATGTAAGCAAAATTAATCC 4872
 Db 1261 GCAAGGATTAAGGAAATTTCTGCTGATTCCTTAATGTAAGCAAAATTAATCC 1320
 Qy 4873 ATTCTTGGGATGAATGTTGCGGGAACCAATGCGGTTCTTTCTACGCGGTAGAGCT 4932
 Db 1321 ATTCTTGGGATGAATGTTGCGGGAACCAATGCGGTTCTTTCTACGCGGTAGAGCT 1380
 Qy 4933 GCTTTGCGAATTAATCTTCTATTAATACTCTGCTTTGATTACGGAACGAAGTA 4992
 Db 1381 GCTTTGCGAATTAATCTTCTATTAATACTCTGCTTTGATTACGGAACGAAGTA 1440
 Qy 4993 AATCCTTTAGTGAAGAAATTAACAAGTCAATGTAACAAGCTTGAATTAATTCATATT 5052
 Db 1441 AATCCTTTAGTGAAGAAATTAACAAGTCAATGTAACAAGCTTGAATTAATTCATATT 1500
 Qy 5053 ACAACGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGAATCGAGAAATGTA 5112
 Db 1501 ACAACGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGAATCGAGAAATGTA 1560
 Qy 5113 TCTGCAATCGTGTCTGATGAACGGAAGCTTTAGTGAATTCGAGTTGAAGA 5172
 Db 1561 TCTGCAATCGTGTCTGATGAACGGAAGCTTTAGTGAATTCGAGTTGAAGA 1620
 Qy 5173 GTAAGTCTTTCAATGTAGATGCAAAAGATCAAAATAACAATTAATTCGCGGAAT 5232
 Db 1621 GTAAGTCTTTCAATGTAGATGCAAAAGATCAAAATAACAATTAATTCGCGGAAT 1680
 Qy 5233 GCAAAATGAGGAAAAAGCGGCTGAGTGTGAGCAACGTTGCTCATCAAAATATTGAAA 5292
 Db 1681 GCAAAATGAGGAAAAAGCGGCTGAGTGTGAGCAACGTTGCTCATCAAAATATTGAAA 1740
 Qy 5293 CAATCAGTTATAGCTATTGTAAATAAGTAATAATTAACAAGCGCAATGATCAATGA 5352
 Db 1741 CAATCAGTTATAGCTATTGTAAATAAGTAATAATTAACAAGCGCAATGATCAATGA 1800
 Qy 5353 AAAAATATCAATGTATGCAAAAGATTAATTAATGACCAATTAATGAGTCCGAGTT 5412
 Db 1801 AAAAATATCAATGTATGCAAAAGATTAATTAATGACCAATTAATGAGTCCGAGTT 1860
 Qy 5413 GAGAGGCAAAAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTTGAATTAAGACA 5472
 Db 1861 GAGAGGCAAAAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTTGAATTAAGACA 1920
 Qy 5473 GTTCTTCTCATGTGATCAAACTGATTTGAACAAGATTTAGAGGAAATAATGA 5532
 Db 1921 GTTCTTCTCATGTGATCAAACTGATTTGAACAAGATTTAGAGGAAATAATGA 1980
 Qy 5533 AATTAAGAAAGGCAATGTTAATGTTCTAGCTGAAATAGCAATGATCAATGAT 5592
 Db 1981 AATTAAGAAAGGCAATGTTAATGTTCTAGCTGAAATAGCAATGATCAATGAT 2040
 Qy 5593 GCGAGAGCTTTCGAGAGCAATGAGCAAGCTGAGTGAAGCTGAGTGAAGTAAAT 5652
 Db 2041 GCGAGAGCTTTCGAGAGCAATGAGCAAGCTGAGTGAAGCTGAGTGAAGTAAAT 2100
 Qy 5653 AAAATTACCAAAATTAATCTGCAATTAATTAATTAATTAATTAATTAATTAATTA 5693
 Db 2101 AAAATTACCAAAATTAATCTGCAATTAATTAATTAATTAATTAATTAATTAATTA 2141

RESULT 6
 AA167644
 ID AA167644 standard; DNA; 1887 BP.
 AC AA167644;
 XX
 XX 27-FEB-2002 (first entry)
 XX
 DE F. necrophorum truncated leukotoxin protein (SH) encoding DNA.
 XX
 KM Leukotoxin; infection; immunisation; liver abscess; foot rot;

QY 7124 TAAAGGCAAAAGGAGAAAATATTTAAGCAACAGCAAAATGAGAAATCGAGGAG 7183
 DB 1501 TAAAGGCAAAAGGAGAAAATATTTAAGCAACAGCAAAATGAGAAATCGAGGAG 1560
 QY 7184 CAGCAATCGAGCAATGCTTGTGTAATTAATTTTGAACAGCTGTAGAGATGAGAAAA 7243
 DB 1561 CAGCAATCGAGCAATGCTTGTGTAATTAATTTTGAACAGCTGTAGAGATGAGAAAA 1620
 QY 7244 ATTCTGAAGAAAGAGAAAGAAAGTTTAAATCTTTAGACGAGTTAACAAAGAACAG 7303
 DB 1621 ATTCTGAAGAAAGAGAAAGAAAGTTTAAATCTTTAGACGAGTTAACAAAGAACAG 1680
 QY 7304 ATTAATAAAGTAATATGATGTAAGAAAAAAATCTTACAACTACAGATATTTTACAGAG 7363
 DB 1681 ATTAATAAAGTAATATGATGTAAGAAAAAAATCTTACAACTACAGATATTTTACAGAG 1740
 QY 7364 ATACTTCTGTAAAGCGATAGAGAGATACACAGGAGAGAAATTAAGCCATTGTGA 7423
 DB 1741 ATACTTCTGTAAAGCGATAGAGAGATACACAGGAGAGAAATTAAGCCATTGTGA 1800
 QY 7424 AGACTTCTGTATATTAATTTGAAAAAATGTAGATATTACACAGAGACAAAGAAATATATCA 7483
 DB 1801 AGACTTCTGTATATTAATTTGAAAAAATGTAGATATTACACAGAGACAAAGAAATATATCA 1860
 QY 7484 CTTCTACTGTGTGTTGGGAACTGCAG 7510
 DB 1861 CTTCTACTGTGTGTTGGGAACTGCAG 1887

RESULT 7

AA167641
 ID AA167641 standard, DNA; 1130 BP.

AA167641;

27-FEB-2002 (first entry)

F. necrophorum truncated leukotoxin protein (BSBSE) encoding DNA.

Leukotoxin; infection; immunisation; liver abscess; foot rot;

bactericidal; vaccine; truncated; BSBSE; de.

Fusobacterium necrophorum.

Key Location/Qualifiers

FT CDS 1..1110

FT /tag= a

FT /product= "truncated leukotoxin (BSBSE)"

XX WO200180886-A2.

XX 01-NOV-2001.

XX 25-APR-2001; 2001WO-US013240.

XX 25-APR-2000; 2000US-00558257.

XX 24-APR-2001; 2001US-00841786.

XX (UNIV) UNIV KANSAS STATE RES FOUNDED.

XX Nagaraia TG, Stewart GC, Narayanan SK, Chengappa MM;

XX WPI; 2002-049245/06.

XX P-PSDB; AAG66006.

PT Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an animal against an infection e.g. foot rot, or liver abscesses caused by the bacterium.

XX Claim 10; Page 95; 108pp; English.

XX The invention provides an isolated Fusobacterium necrophorum leukotoxin

CC polypeptide and its truncated versions. The leukotoxin protein can be
 CC expressed by standard recombinant methodology. The leukotoxin gene is
 CC useful for preparing a vaccine which confers effective immunity against
 CC infection caused by F. necrophorum. The vaccine is useful for immunising
 CC an animal against liver abscesses caused by F. necrophorum and for
 CC preventing foot rot caused by F. necrophorum infection. The present
 CC sequence represents a DNA encoding a F. necrophorum truncated leukotoxin
 CC protein (BSBSE)

XX Sequence 1130 BP; 444 A; 135 C; 234 G; 317 T; 0 U; 0 Other;

Query Match 11.6%; Score 1130; DB 6; Length 1130;

Best Local Similarity 100.0%; Pred. No. 9.1e-210;

Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGGCATCAAAATTAAGTTAGAGACAAAGAAAGAGATTCAGATCTTAAAAA 60
 DB 1 ATGAGCGGCATCAAAATTAAGTTAGAGACAAAGAAAGAGATTCAGATCTTAAAAA 60
 QY 61 GTTTAATGATTTGGGATTTGATTAACATATGACGGTGAAGGCTATGATCAATC 120
 DB 61 GTTTAATGATTTGGGATTTGATTAACATATGACGGTGAAGGCTATGATCAATC 120
 QY 121 ACCGCGACTGAGAAATTTTGAACAAAAATAGAAAAAAGATATGTTATGACATTA 180
 DB 121 ACCGCGACTGAGAAATTTTGAACAAAAATAGAAAAAAGATATGTTATGACATTA 180
 QY 181 ACAACCAAGATTCAAGGGGAGAACGCTTTTAACATTTTAATAGATTTGCTTTACAG 240
 DB 181 ACAACCAAGATTCAAGGGGAGAACGCTTTTAACATTTTAATAGATTTGCTTTACAG 240
 QY 241 AATAATATAGCAAAATCTATATTTTGGGAAAAAGATATGACGGGATTAATCTTTT 300
 DB 241 AATAATATAGCAAAATCTATATTTTGGGAAAAAGATATGACGGGATTAATCTTTT 300
 QY 301 AACTTGTCAATGAAAAATTTGAAGTAGATGATATCAACGAAATTCAGAAAAATAA 360
 DB 301 AACTTGTCAATGAAAAATTTGAAGTAGATGATATCAACGAAATTCAGAAAAATAA 360
 QY 361 ATTGAGCAAAATTTATATTTCTTAAGCTCGGAAGGATGCGATGAGAAAAATGAGAT 420
 DB 361 ATTGAGCAAAATTTATATTTCTTAAGCTCGGAAGGATGCGATGAGAAAAATGAGAT 420
 QY 421 ATCAATGCTGCTCTTTTCAATCTATTTTCCAAAAAAGATGATTTAAGAGGCTTG 480
 DB 421 ATCAATGCTGCTCTTTTCAATCTATTTTCCAAAAAAGATGATTTAAGAGGCTTG 480
 QY 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGATATTCAGATGAGAAAAATGAGAT 540
 DB 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGATATTCAGATGAGAAAAATGAGAT 540
 QY 541 ATTCATTTGAATCCGAATGAGAAAGATTCAGATGAGAAAAATGAGATGAGAT 600
 DB 541 ATTCATTTGAATCCGAATGAGAAAGATTCAGATGAGAAAAATGAGATGAGAT 600
 QY 601 ATCGGTTTATGCGCGGATTTAGATGAGAAAGATTCAGATGAGAAAAATGAGAT 660
 DB 601 ATCGGTTTATGCGCGGATTTAGATGAGAAAGATTCAGATGAGAAAAATGAGAT 660
 QY 661 ACAGATTTTAAATTTAGTCAATATTTAGTCAATATTTAGTCAATATTTAGTCAAT 720
 DB 661 ACAGATTTTAAATTTAGTCAATATTTAGTCAATATTTAGTCAATATTTAGTCAAT 720
 QY 721 TTTAAAGCTACAGCAAAATCTGAGATATTTATTTCTTACAGCTACATATTTCTCT 780
 DB 721 TTTAAAGCTACAGCAAAATCTGAGATATTTATTTCTTACAGCTACATATTTCTCT 780
 QY 781 CAAAAAGCTATGGAAGAAAAATCACTGTTGGAAGAGATGAGAAATATGTTAAAGAT 840
 DB 781 CAAAAAGCTATGGAAGAAAAATCACTGTTGGAAGAGATGAGAAATATGTTAAAGAT 840
 QY 841 AATACCAAGCAAAATATGATCTGATGATTTGGAAGAGATGAGAAATATTTAAATTT 900

Db 841 AATACCAAGCAATATGTAATCTGATGCTGTAATTGGAGCAGATGAAATATTAATAATT 900
Qy 901 AGTGGAAAAGCTACAAATGGAGATTATTAAGAAAAGGAGAAAAGAACTTATTAAC 960
Db 901 ACTGGAAAAGCTACAAATGGAGATTATTAAGAAAAGGAGAAAAGAACTTATTAAC 960
Qy 961 ACTCCTTTAAGTTTATCAGATGTGGAACTTCGTAAGTAATTAAGAAAAGTCATA 1020
Db 961 ACTCCTTTAAGTTTATCAGATGTGGAACTTCGTAAGTAATTAAGAAAAGTCATA 1020
Qy 1021 GGAAGAAAGTTGACATTCAGCTGAGCAAGAAATTTCTATGAGCACTTATTAAGT 1080
Db 1021 GGAAGAAAGTTGACATTCAGCTGAGCAAGAAATTTCTATGAGCACTTATTAAGT 1080
Qy 1081 AAGCTTGCAAGACACTCTTTTAGCTTTGTTACAGGTTCTATTTCTCTAT 1130
Db 1081 AAGCTTGCAAGACACTCTTTTAGCTTTGTTACAGGTTCTATTTCTCTAT 1130
RESULT 8
AAT78868
ID AAT78868 standard; cDNA; 5361 BP.
XX
AC AAT78868;
XX
DT 08-OCT-1997 (first entry)
XX
DE P. falciparum liver stage antigen-3 coding sequence.
XX
KM Plasmodium falciparum; pre-erythrocyte; liver stage antigen; serum;
KM prophylaxis; Thai strain; gene organization; exon; intron; hydrophobic;
KM glycosyl-phosphatidylinositol membrane anchoring sequence; antibody;
KM vaccine; immunotherapy; malaria; ds.
XX
OS Plasmodium falciparum.
XX
FN WO9641877-A2.
XX
PD 27-DEC-1996.
XX
PF 12-JUN-1996; 96WO-FR000894.
XX
PR 13-JUN-1995; 95FR-00007007.
XX
PA (INSP) INST PASTEUR.
XX
PI Drullhe P, Daubersies P;
XX
DR WPI; 1997-065464/06.
XX
DR P-PSDB; AAM24790.
XX
PT Plasmodium falciparum poly-peptide(s) and related nucleic acids - derived
PT from the liver stage antigen-3, useful for malaria vaccine prodn. and
PT diagnosis.
XX
PS Claim 20; Fig 2A-I; 69pp; French.
XX
CC This sequence corresponds to the coding sequence for a Plasmodium
CC falciparum strain KI pre-erythrocytic liver stage antigen-3 (LSA-3)
CC protein. The gene sequence was isolated by screening a P. falciparum
CC strain T9/96 library with serum from a missionary treated by prophylaxis
CC (for strain T9/96 see FR9101286). Of 20 clones isolated, clone 7295 was
CC used to screen a library generated from Thai strain KI. One clone
CC contained a 6.85 kb insert including the genomic sequence AAT78867. The
CC gene comprises a 1.8 kb region encoding 3 major blocks of 4 amino acid
CC repeats and a 3' hydrophobic region corresponding to a glycosyl-
CC phosphatidylinositol membrane anchoring sequence. The invention relates
CC to new polypeptides of at least 10 amino acids derived from the LSA-3
CC polypeptide with the exception of the peptides AAM24791-4. The LSA-3
CC peptides can be used to raise antibodies and as vaccines for
CC immunotherapy of malaria
XX
SQ Sequence 5361 BP; 2388 A; 431 C; 1169 G; 1373 T; 0 U; 0 Other;

Query Match 1.2%; Score 116.2; DB 2; Length 5361;
Best Local Similarity 41.7%; Pred. No. 2.5e-12;
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;
Qy 5126 GTTCTGATGAAAACGAGACCTTATAGTGAATTCGAGTTTGAAGAGTAAAGTTCTTCA 5185
Db 773 GTGTGAAGAAAGATGATAGCTTCAAGTGTGTAAGAAATGATTAATCAATGTAAGAAA 832
Qy 5186 ATGTAGATGCAAAAGATGTAAGAAAACAAATTAATACA---TTGCCGGAATTCGAAATGAG 5242
Db 833 ATGTAGCTCCAACTGTGTAAGAAATCGTAGCTCCAAAGTGTGTAAGAAAGTGGCTCCAA 892
Qy 5243 GAAAGCCGCTGAGATTGAGACCAACAGTTGCTCATCAATATTTGAAACAAATCAGTTA 5302
Db 893 GTGTGAAGAAAGATGTAAGAAATGTTGAAGAAAGTGAAGTGAAGAAATGTTGAAGAAA 952
Qy 5303 TAGCTATGTAAGAAAACATGAATTAACACGCGGAATGATCAAGATGAAAAAATATCA 5362
Db 953 GTGTAGCTGAAAATGTTGAAGAAAGTGAAGTGAAGAAATGTTGAAGAAAGTGAAGTGA 1012
Qy 5363 ATGTAGCTGCAAAAGATTAATCTATGACCAATATCTATGACAGCTCGAGATTGAGAGCAA 5422
Db 1013 ATGTGAAGAAATGTAAGCTCCAACTGTTGAAGAAATGTAAGCTCCAACTGTTGAAGAAA 1072
Qy 5423 AAGGACCTGTGTCAGCAAGACCTTCGCAAGTACTACTGAAATTAAGACAGTTCTTCTC 5482
Db 1073 TTGTAGCTCCAAAGTGTGTAAGAAAGTGGGCTCCAAAGTGTGTAAGAAAGTGAAGAAA 1132
Qy 5483 ATGTGATCAAACTGATATTAATGACAAAGATTAGAGGAAGAAATTAAGAAATTAAGAAA 5542
Db 1133 ATGTGAAGAAAGTGAAGTGAAGAAATGTTGAAGAAAGTGAAGTGAAGAAATGTTGAAGAAA 1192
Qy 5543 AGCGAAATGTAAGTCTAGCTGTAAGAAATAGAGTCAAGTGCACAAATCGACAGTGC 5602
Db 1193 GTGTAGCTGAAAATGTTGAAGAAAGTGAAGTGAAGAAATGTTGAAGAAAGTGAAGTGA 1252
Qy 5603 TTTCGAGACCAAGTGAAGCAAGCTGAGAGAGCTGAGTGAAGTGAATTAATTAATTAAC 5662
Db 1253 ATGTGAAGAAATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGAAA 1312
Qy 5663 AAAATCTTGCACATATTAAGAAATTAAGTACTCAAAATGTAAGAAATGCTTTGTAAGAAA 5722
Db 1313 TTGTAGCTCCAAAGTGTGTAAGAAAGTGGCTCCAAAGTGTGTAAGAAAGTGTAGAGAAA 1372
Qy 5723 GC---AAATCTCATCTATCTATTAAGAAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5779
Db 1373 ATGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGAAA 1432
Qy 5780 GAGCTGAGTGAAGCAAGTCTGTAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5839
Db 1433 GTGTAGCTGAAAATGTTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGAAA 1492
Qy 5840 TAAATCATGCAAAATATCATGCGAAGAAATGTCGAGTATTAACAGATGTAAGTCGCG 5899
Db 1493 ATGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGAAA 1552
Qy 5900 TAAATTCCTAATTAATGAGCAAGCAAGTGTGAGAGTGGCCCGGAGCAATATAGAGCTCAA 5959
Db 1553 TGTAGCTCCAACTGTGTAAGAAATCGTAGCTCCAACTGTGTAAGAAATGTTAGTCCAA 1612
Qy 5960 CCAAGTGAATGAATTAACAGATCTCAAGATCAATGAAGATTAAGATTAAGATTAAGATTAAG 6019
Db 1613 GTGTGTAAGAAAGTGTGCTCCAAAGTGTGTAAGAAAGTGAAGAAATGTTGAAGAAA 1672
Qy 6020 CTAAGAGAAAGATGATATTAATTAATCAAGGCAAGTGAAGTGAAGTGAAGTGAAGTGA 6079
Db 1673 GTGTAGCTGAAAATGTTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGAAA 1732
Qy 6080 AAGATTTCAAAATCTTAATTAATTAACGAAGCTTAATCAAGAAAGAAAGAAATTAAGTAA 6139
Db 1733 ATGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1786


```

Db 3936 CAAGAATCGAAAAAGATATCTGATTTAAAGATCTTGAGAGATATATTAAAG----- 3988
Qy 8328 TTCTGTAGATCAGAAAAATGAGAGAAATTCGCCGAGAGAGATGAAATCTTCTACAGC 8387
Db 3989 --AAGTAAAGAAATCAAAAGAACTTGAAAGTAAATTTTGAAGATTTTAAAGATTAAA 4046
Qy 8388 ACAAGTAAATCTGAATCCGAGTTCGTTACGAAAGCAAGATTTAGAAAATATGATTA 8447
Db 4047 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATAGAAAAAGATCATTTTGAAGA 4106
Qy 8448 CACAAAAAATATATTTCGAAGTCATGCTCTTCTTTAAATGATCAAGAAATGAAGC 8507
Db 4107 ATTCGAAGAGAGAGCTGAGAAATATAAATCTTGAAAGCATATATTAAAGAAAGTATC 4166
Qy 8508 GAATATAGATCTTTAGCCGTCGTCATGCACAGAGAAACAAACAAAGCATTTAC 8567
Db 4167 TTCATTAGAGTTGAGAAAGAAAAAATTAGAAGAGTACGAAATTTAAAGAAAGAGT 4226
Qy 8568 GAGATCAACAGATTACTTCTCAACCTGTAAATGAGAGAAACGTAATCTCACTTCGTC 8627
Db 4227 AGAATCATATATTAAGTGTGATGCGCATATAAAGTTTGAAGAAAGATGATTTAGAGA 4286
Qy 8628 AAAACCTTGCCTTAAATGAAATTTATGGAATGTAAAGAACTGAGAGCCTTACT 8687
Db 4287 AGTAAATGA-TTTAAAGAGAGATATATTAAGACATGTTAAAGGAGATATGA--ATTAGG 4343
Qy 8688 CGAGCCGAGAAACAGACGCCGTTGAAAAATTATCAAAAGATCTACAGAGCATTTGCTGC 8747
Db 4344 GGATATGATTAAGAGAAAGTTTGAAGATGTAACACAAACCTTGAGAGAAAGTTGATC 4403
Qy 8748 AGGAAATTGGAAATTGAGATTAATAGAAACGATTGCAAGATTAATCAGATTGTAG 8807
Db 4404 CTTAAAGATGTTTATCTAGTGCATTAGCATGATGAGAAACAAATGAAAAACAAGAA 4463
Qy 8808 AGTCAACGAGAGCGAACCAGAGGCTTTCGGAAGAAATGATATTTCTGTGAAAAA 8867
Db 4464 AAAAGCTCAAGACCTTAAAGTGAAGATATTTAAAGAGGTTTAAAGAAAGAAC 4523
Qy 8868 TACAATTTTCAGGGGAAACAAATTCATCTTGAAGATTAAGCCAGAAATTTGTGAACGG 8927
Db 4524 AAAAGAAAAAATACAAAAAAGAAAGTAAAGTTTATTAAGATTAAGAAACAAAGAA 4583
Qy 8928 AAGGTAAATGTAGATGCTTTGATGAATGAACCTTGATGATCTCAAGAGAAAAAGTGTG 8987
Db 4584 TGAATATGTAAGATTGAATGAAAGATGAAGATTAAGAAAGAAATGTAGAAAGATAT 4643
Qy 8988 CTATGTGGAATTTGATTTGAAATGTTGATTAATATGATTAAGAAAAATGTAGA 9047
Db 4644 AGAAGAGATATAGAAAGAAAGATTAAGTTGAAGATTAAGATGAAGATAT 4703
Qy 9048 AGCCAAATCGGAGACATGCTATTGTGAAGAACTACTGGAAGAAACAAATATCAAGCAT 9107
Db 4704 AGGTGAAGCAAGATGAGATTTATGATTAATAGTCCAAAGAAAGAAAGCATTTGAAAA 4763
Qy 9108 TACAAGAGCAAAAGTAAATATTTCTTGAAAGAGAGACGCTGACGTCAGCTCAATATC 9167
Db 4764 GGTAAAGCGAAAAAGAAAAAATTAGAAAAAAGTTGAAGAGGTTAGTGCTTAA 4823
Qy 9168 GAATGTACATTTCCATGAGATGATTTTAAAAATTTGGCAAGACGATATGCTTC 9227
Db 4824 AAAACGCTAGACGAAGTAAATATATGTTCAAAAAATGTATTAAGAAAGTTGATTAAGA 4883
Qy 9228 TCAATTAATTAACCAAAAAATTCAAAAAATAATATTACTTAGCATCA 9274
Db 4884 AGTATCTTAAGCTTTAGAAATCAAAAAATGATGTTACTAATGTTTAA 4930

```

RESULT 9
 ABK50883
 ID ABK50883 standard; DNA; 5529 BP.
 AC
 XX ABK50883;
 XX

```

DT 07-AUG-2003 (revised)
DT 30-JUL-2002 (first entry)
XX
XX DNA encoding Plasmodium falciparum liver stage antigen-3 (LSA-3) #1.
DE
XX
XX Liver stage antigen-3; LSA-3; vaccine, Th1-inducing adjuvant;
KM malaria parasite; malaria; protein-specific cytotoxic T cell response;
KM gamma-interferon; gene; de.
XX
XX Plasmodium falciparum.
OS
XX
XX Key Location/Qualifiers
FH 1..5529
FT CDS
FT /*tag= a
FT /product= "LSA-3"
FT /note= "Liver stage antigen-3"
FT /transl_except= (pos:5524..5526, aa:Lys-Glx)
FT 1..198
FT exon /*tag= b
FT /number= 1
FT intron 199..366
FT /*tag= c
FT /number= 1
FT exon 367..5529
FT /*tag= d
FT /number= 2
FT
PN
XX
XX EPI201250-A1.
PD 02-MAY-2002.
XX
XX 25-OCT-2000; 2000EP-00203724.
PF
XX 25-OCT-2000; 2000EP-00203724.
PR
XX 25-OCT-2000; 2000EP-00203724.
XX
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA (INSP ) INST PASTEUR.
XX
XX Cohen J, Drullhe P;
PI WPI; 2002-373883/41.
DR
XX P-PSDB; AAU96699.
DR
XX
XX Vaccine for treatment or prevention of malaria, comprises liver stage
PT antigen and adjuvant that induces Th1 response.
XX
XX Disclosure; Page 16; 56pp; English.
PS
XX
XX The invention describes a vaccine comprising a Th1-inducing adjuvant (I)
CC and a protective liver-stage antigen (LSA), or its immunogenic fragment,
CC from a human malaria parasite, where LSA is a fragment of LSA-3, then (I)
CC is not montanide. The vaccine is useful for treating or preventing
CC malaria, specifically where caused by Plasmodium falciparum. The vaccine
CC is formulated in SBA52 adjuvant, an oil-in-water emulsion containing QS21
CC and 3D-MPL (de-O-acylated monophosphoryl-lipid A, the combination of
CC which results in strong induction of a circumsporozoite protein-specific
CC cytotoxic T cell response, which is not generally induced by vaccines
CC based on recombinant proteins, and synergistically increases production
CC of gamma-interferon. This sequence encodes the liver-stage antigen-3 (LAS
CC -3) from the Plasmodium falciparum strain KI used to develop the
CC recombinant proteins for the vaccine. (Updated on 07-AUG-2003 to correct
CC OS field.)
XX
XX
SQ Sequence 5529 BP; 2473 A; 435 C; 1174 G; 1447 T; 0 U; 0 Other;

```

Query Match 1.2%; Score 116.2; DB 6; Length 5529;
 Best Local Similarity 41.7%; Pred. No. 2.5e-12;
 Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;
 Qy 5126 GTTCTGATGAACGAGCTTATGATGATTTCTGAGATTGAAGAGTAACTTCTTCA 5185
 Db 941 GTGTTGAAGAAAGTATGAGCTTCAAGTGTGATGAAGATTAATCAAGTATGAAGAAA 1000

Db 3152 AAGTTTGAAGAAACAATGATGATTAAGCATTTTGTAGTAAATATTGTATTAATGTA 3211
 Qy 7391 ATACTCAGGAGAGAGAAATTAAGCCATTGTGAAGCTTCTGATATTATTTGAGAAAAATG 7450
 Db 3212 AAGGAATACAGAAATTTATTACAGATG-----TTTGGAGTATAGAAACCACTA 3265
 Qy 7451 TTAGATTTCACACAGAGACAGAAATATATCACTTACTGCTGTTGGAACTCGAG 7510
 Db 3266 TAGATATCCATCAGAAAGAAAGTTGATTGAATGAATAATGTGTTAGTTTGA 3325
 Qy 7511 GCTTGCTCCGCAATCAGAGACAGTGGCAGTTACAAATATTAAGAAATCCGAGTGA 7570
 Db 3326 ATATATATGAAATATGAAAGAGGTTATTAATTAATTAAGAAATATTTCAAGTACTG 3385
 Qy 7571 CTGTTGAAATTCCTTTGTGAAGCAGCTGAAAAAGTAATGTAGATCGATAT----- 7625
 Db 3386 AAGGTGTTCAAGAACTGTATCTGAACATGTAGAACAAATTTATATGTGATGTTGATG 3445
 Qy 7626 ---TACAGAAATGTTGCTTTAACAGCATTCAGAGTCTGTAGAGCAATTGGAAATAG 7681
 Db 3446 TTCTGCTATGAAAGATCAATTTTATAGAAATTTAAATGAGGCAAGAGGTTGAAAGAA 3505
 Qy 7682 GAGCTGCCATGCAAAATTAATTTCTAATGAAAGTCAATATCAATATTAATAATTTCTA 7741
 Db 3506 TGTTTTATTTTGAAGATGATTTTAAAGTGAAGTGAATGATTAATCTGTAGAGAA 3565
 Qy 7742 AGCTATTAGAAAAAATATTGATGTTATTGTAAAAAGATTAATCGAATGAGCGAAG 7801
 Db 3566 TTAAGGATGACCGGTTCAAAAAGAGTGAAGAAAGAACTGTAGTATTTATTTAGAGAA 3625
 Qy 7802 CAAAAGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAGAAATG 7861
 Db 3626 TCGAAGAAATATTGTAGATGTTATAGAGAAAGAAAGAAATTTAACAGCAAGATGA 3685
 Qy 7862 AATGATTCAGAGGTTGAATTTGAGAGAGATTTTCAATGAAGAAATTAAGATACTA 7921
 Db 3686 TTAGATGCAATGAGAAATCCATGAAATATCTTCAGATTTCAAGAGAACTGAAATCTA 3745
 Qy 7922 GCCCTTCTAAAGAAATTTGAGAGAAATCAATGTCAAAGTGAAGAAAGCAAGATGA 7981
 Db 3746 TTAAGATTAAGAAAAAGATGTTCTACTGTTGTGAAGAAATTCAGACCAATGATATGG 3805
 Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGCAGTAGCAGGGGCAAG 8027
 Db 3806 ATGAAAGTGTGAGAAAGTTTATGAATGAATAATGAGAGAGATTAATGAAGATG 3865
 Qy 8028 AATTATTTCCGAGCAAAAGATGCCGAGCTCTTATTGAAAGTATAGTACAAATCCGG 8087
 Db 3866 CTGTGAAATTAATGATCTTACTAGCAAACTTATTTGAAGAACTCAGAGGTTAAATGAG 3925
 Qy 8088 AAGAATATTTTTCATGCAAGATATGTGAATATGAGCAACACATTAATGAAGTAAC 8147
 Db 3926 TTAGAAGCAATTTAATTAAGATATGAGAAAAATTTAAGAAATTTAGAAAAAGC--ATTATC 3983
 Qy 8148 AGCAGTTTCTAAAGCAGTAACAGGTTCTGTATTTGGAGAGGTTGAGTCAACCAAGCAGA 8207
 Db 3984 AGAAGATCTTAAGAAATATATAGATGCAAAAGATGATACATTAGAAAAAGTTATTGAGA 4043
 Qy 8208 ACTTACTGTGACAGTAATACTATGTATGAAGTTGAGAGAGAAATTTGTTTCAAGACAA 8267
 Db 4044 GGAACATGATATTAACGACGCTGTGATAGATGATGATTAATTAAGATGTGAAGAGAA 4103
 Qy 8268 TCGATTGAATGCATTTCTAAAGTAGAAGTTTGTGATGAAGTAAGTAATCTTCAACAC 8327
 Db 4104 CAAGATCGAAAAAGTATCTGATTTAAAGATCTTGAAGAGATATATTTAAAG----- 4156
 Qy 8328 TTCTGTAGTATCAGAAATGAGAGAAATGCCCCAGCAGAGATGAATATCTTCAACAC 8387
 Db 4157 --AAGTAAGAAATCAAGAACTTGAAGTGAATTTTAAGAAATATTAAGAAATTAAG 4214
 Qy 8388 ACAAAATATATCTGAATCCGTAGTCTGTTTGAAGAAAGCAAGTTATGAATAATATGATTA 8447
 Db 4215 AACTATTGAACAGATATTTTGAAGAGAAAAAGAAATAGAAAAAGATCATTTTGAAGAA 4274

Qy 8448 CACAAAAATATATTTTCAGAGTCAATGCTCTTGCTTTAAATGATACAAAGAAATGAGC 8507
 Db 4275 ATTGGAAGAAAGACTGAAAGAAATTAAGAAATCTTGAAGCAATATATTAAGAAAGATATC 4334
 Qy 8508 GAATATAGATCTTTAGCGGTAGCCGGTGTGATGACACAGAAACAAACAAATTTTAC 8567
 Db 4335 TTCAATTAAGATTAAGAAAGAAAAAATTTGAAAGATACAGAAATTTAAAGAAAGAGT 4394
 Qy 8568 GAGATCAAAACAAGTTAATCTTACAACTGTAAATGAGAGAAACGATCTCAACTGCTGC 8627
 Db 4395 AGAACATTTAATTAAGTGTGATGCGCATATTAAGGTTTGAAGAAAGATGATTTAGAAAG 4454
 Qy 8628 AAAAGCTTTGCTAAAAATGAAAAATTAAGAAATGTAAGAAAGAACTGGAGAGGCTTAGT 8687
 Db 4455 AGTAGATGA--TTTAAAGGAAGTATATTAAGCATGTTTAAAGGAGATATGGA--ATTAGG 4511
 Qy 8688 CGAGCGGAAACAGACCGGTTGAAAAATTTATCAAAAGATCTACAGAGCATTTGTTGC 8747
 Db 4512 GGATATGATTAAGGAAAGTTTAGAAGATGTAAACAACAACCTGGAGAAAGATTTGAATC 4571
 Qy 8748 AGAAATTTGGAAATTTGAGATAATTAAGAAACGATTTGCAAGATATATAGATTTGTAG 8807
 Db 4572 CTTAAAGATTTTATCTAGTGAATGAGCATGATGATGAACAAATGAAGAAACAGAA 4631
 Qy 8808 AGTCAACGAGACCGAAACCAAGAGGTCTTGCGAAAGATGATTTCTGTGAAAAA 8867
 Db 4632 AAAAGCTCAAGACCTTAAGTTGAGAAAGATTTATTAAGAAAGAGGTTAAAGAAAGCC 4691
 Qy 8868 TACAATTTCAAGGGAACAAATATCATCTTGAAGATTAAGCCAGATTTGTTGAAACCG 8927
 Db 4692 AAAGAAAAAATTAACAAGAAAGAAAGTAAAGTTGATTTAAGATTAAGAAACCAAGAA 4751
 Qy 8928 AAGGTAAATGTAGATGCTTGAATGAACCTTAGATGATCTACAGGAAAAAGTGGCG 8987
 Db 4752 TGAATTAAGTAAAGTTGAATGAAGATGAAGATTAAGAAAGATGTAGAAAGATAT 4811
 Qy 8988 CTATGTGGAATTTGATTTGAAAAATGTTGATGTAAATATGTGATTAAGAAAAATGTGA 9047
 Db 4812 AGAAGAAATATTAAGAAAGATTAAGTTGAAGATTAAGTAAAGATTAAGTAAAGATAT 4871
 Qy 9048 AGCCAAATTCGAAAGACATGCTATTTGTGAACATCTAGGAAACAAAGAAATTCAGACTT 9107
 Db 4872 AGGTGAAGACAAAGATGAAGTTATATAGTCCAAAGAAAGAAACCATTTGAAAA 4931
 Qy 9108 TACAAGCAAAAGTAATATTTCTTGAAAAAGAGACGCTGACGCTGCAATATC 9167
 Db 4932 GGTAAAGCGAAAAAGAAAAATTAAGAAAAAGTTGAAGAGGTGTTAGTGTCTTA 4991
 Qy 9168 GAATGTACACATTTCCATGAGATGATTAATTAATTTGCAAAAGCAGTATGATCTTC 9227
 Db 4992 AAAACAGTAAAGCAAGTAAGAAATATGTTCAAAAAATTTGATTAAGAAAGTTGATTAAGA 5051
 Qy 9228 TCAATTAATTAACAAAAATTCAAAAATATATATTACTTTAGCATCA 9274
 Db 5052 AGTATCTTAAGCTTTAGATCAAAAAATGATGTTACTATGTTTAA 5098

RESULT 10
 AAT78867
 ID AAT78867 standard; DNA; 6152 BP.
 XX
 AC AAT78867;
 XX
 DT 08-OCT-1997 (first entry)
 XX
 DE P. falciparum liver stage antigen-3 genomic sequence.
 XX
 KW Plasmodium falciparum; pre-erythrocyte; liver stage antigen; serum;
 KW prophylaxis; Thai strain; gene organisation; exon; intron; hydrophobic;
 KW glycosyl-phosphatidylinositol membrane anchoring sequence; antibody;
 KW vaccine; immunotherapy; malaria; ds.

OS Plasmodium falciparum.
XX Key Location/Qualifiers
XX CDS 77..5605
FT /tag= a
FT /product= "liver stage antigen-3"
FT exon 77..274
FT /tag= b
FT /number= 1
FT Intron 275..442
FT /tag= c
FT /number= 1
FT exon 443..5602
FT /tag= d
FT /number= 2
XX MO9641877-A2.
XX 27-DEC-1996.
XX 12-JUN-1996; 96WO-FR000894.
XX 13-JUN-1995; 95FR-00007007.
XX (INSP) INST PASTEUR.
XX Drulhe P, Daubersies P;
XX WPI; 1997-065464/06.
XX Plasmodium falciparum poly:peptide(s) and related nucleic acids - derived
XX from the liver stage antigen-3, useful for malaria vaccine prodn. and
XX diagnosis.
XX Claim 20; Fig 1A-E; 69pp; French.
XX
XX This sequence corresponds to the genomic sequence encoding a Plasmodium
XX falciparum strain KI pre-erythrocytic liver stage antigen-3 (USA-3)
XX protein (AAW24790). The gene sequence was isolated by screening a P.
XX falciparum strain T9/96 library with the serum from a misionary treated
XX by prophylaxis (for strain T6/96 see PR9101286). Of 20 clones isolated,
XX clone 7295 was used to screen a library generated from Thai strain KI.
XX One clone contained a 6.85 kb insert including the sequence presented
XX here. The gene organization comprises a first exon, a short 168 bp intron
XX and a 5 kb second exon containing a 1.8 kb region encoding 7 blocks of 4
XX amino acid repeats and a 3' hydrophobic region corresponding to a
XX glycosyl-phosphatidylinositol membrane anchoring sequence. The invention
XX relates to new polypeptides of at least 10 amino acids derived from the
XX USA-3 polypeptide with the exception of peptides AAW24791-4. The USA-3
XX peptides can be used to raise antibodies and as vaccines for
XX immunotherapy of malaria
XX
XX Sequence 6152 BP; 2725 A; 463 C; 1215 G; 1749 T; 0 U; 0 Other;
SO
Query Match 1.2%; Score 116.2; DB 2; Length 6152;
Best Local Similarity 41.7%; Pred. No. 2.6e-12;
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;
QY 5126 GTTCTGATGAACGGAAGCTTGTAGTCTGATCTGATGAAGAGTAAGTTCTTCA 5185
DB 1017 GTGTTGAAGAAAGTATAGCTTCAAGTGTATGAAAGTATAGATTCAGATTGAGAAA 1076
QY 5186 ATGATGATGCAAAAAGATCAAAAAACAATAATACAA---TTGCCGGAATGCAAAATGAG 5242
DB 1077 ATGATGATGCAAACTGTTGAGAAATCGTACTCCAAAGTGTGTAAGAAAGTGGCTCAA 1136
QY 5243 GAAAAAGCGCTGAGTGTGAGCAACAGTGTCTATACAAATATTGAAAAACAATCAGTTA 5302
DB 1137 GTGTTGAAGAAAGTATGAAAGAAATGTTGAAGAAAGTATGAGCTGAAAGAAAGTTGAAGAAA 1196
QY 5303 TAGCTATTGTAATAAACAAGTAATAATTACACGCGGAGATGATCAAGATAGAAAAATATCA 5362
DB 1197 GTGATGCTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAA 1256

QY 5363 ATGTGATGCAAAAAGATTAATCTATGACCAAAATCTATAGCAGTGGAGTGTGAGAGCAA 5422
DB 1257 ATGTTGAAGAAATCTAGCTCCAACTGTTGAAGAAATGTAGCTCCAACTGTTGAAGAAA 1316
QY 5423 AAGAGACCTCTGTGCAAGAGCTTCTGCAAGTACTTCACTTGAATTAAGACAGTTCTCTC 5482
DB 1317 TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAA 1376
QY 5483 ATGTTGATCAAACTGATATTGCAAAAGATTAGAGAAAGAAAATTAAGAAATTAAGAAA 5542
DB 1377 ATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA 1436
QY 5543 AGCAAAAGTTATATGTTCTAGCTGAAAATACAGTCAAGTGTCAAAATGCGACAGTGC 5602
DB 1437 GTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1496
QY 5603 TTTCCGAGCAAGTGAACAGCTGAGTGAAGTGAAGTGAAGTGAATTAATAATTAACAC 5662
DB 1497 ATGTTGAAGAAATCTAGCTCCAACTGTTGAAGAAATGCTAGCTCCAACTGTTGAAGAAA 1556
QY 5663 AAAATACCTTCTGACATATTAATAAATAGTACTCAAAATGTACGAAATGCTTTGTTAATA 5722
DB 1557 TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAGTGTGAAGAAAGTGTGAAGAAA 1616
QY 5723 GC---AAATCTATTCATTAATTAATAAACAATTGGAATGAGCTGGAAGTGAAGCTGAG 5779
DB 1617 ATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA 1676
QY 5780 GAGCTGAGTGAACAGTCTGTAGCAGTGAATTAAGTGAATTAATTAATTAATAGCAGAAAT 5839
DB 1677 GTGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTGAAGAAA 1736
QY 5840 TAAATCATGCAAAAAATCACTGCGAAGGAATGTGGAAGTATTAACAGAGCTGATCCG 5899
DB 1737 ATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA 1796
QY 5900 TAAATCTTAATTAATGACAGAAACAGTGTGAGAGTGGCCGTCAGCAATTAAGAGCTTCAA 5959
DB 1797 TGTAGCTCCAACTGTTGAAGAAATGTAAGTGTGCTCCAACTGTTGAAGAAATGTAAGTGTGCTCCAA 1856
QY 5960 CCAAGTGAATGAATTAACAGATCTACAAAGATCAAAAGATGTAATAAGATTCTACAGTATTTG 6019
DB 1857 GTGTTGAAAGAAAGTGTGCTCCAAAGTGTGAAAGAAAGTGTGAAGAAATGTTGAAGAAA 1916
QY 6020 CTAAAGAGAAACAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6079
DB 1917 GTGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1976
QY 6080 AAGTATTCAAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6139
DB 1977 ATGTTGAAGAAAGTGTAGCT-----GAAAATGTTGAAGAAATGTAAGTGTGCTCCAACTGTTG 2030
QY 6140 AAAAAGATTTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6199
DB 2031 AAGAAATCGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAAAGTGTGTA---GAAAGT 2087
QY 6200 CCGCTGTTTCAAGACAGCGAGTGTGAGAGTGTGAGTGTGATTAATTAATTAATTAATTAATTA 6259
DB 2088 TGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAATG 2147
QY 6260 AAACAGAGCTCTGTGAAGAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6319
DB 2148 TTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATG 2207
QY 6320 GAGATTACGCAATTAATCGAGTGTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 6379
DB 2208 TAGCTCAACTGTTGAAGAAATGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAAAGT 2267
QY 6380 TAGAGCTTCTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6436
DB 2268 TTGTGAAGAAAGTGTGCTCCAAAGTGTGAGAAAGTGTGAAGAAATGTTGAAGAAAGTGT 2327

QY 6437 AAACCTCAAACTGCTGATGAAGGTTCCGAGAAAGAGCTGAAATTCACAGAGATCTTAAGC 6496
DB 2338 TAGCTGAAAATGTTGAAAGAAAGTGTAGCTGAAAATGTTGAAAGAAAGTGTAGCTGAAAATG 2387
QY 6497 AAGGAATTTCTCTTTTGGAGTCGAGTCGACAGCCGCGGTAGAGCCGAGAGTGGCAG 6556
DB 2388 TTGAAGAAAGTGTAGCTGAAAATGTTGAAAGAAATGCTAGCTCCAACTGTTGAAGAAATG 2447
QY 6557 GAACCGTTCCGTTAAATCAATTTGACAGAAAGACGGAAGTAGATGTGAAGAAAGAGA 6616
DB 2448 TAGCTCCAACTGTTGAAGAAATGTTGAGCTCCAAAGTGTGTAGAAAGTGGCTCCAAAGT 2507
QY 6617 TTTTGGTAAAAAGCTGAGATTACAGCAAAACGTTAATGTTCTGTTGCAATTTGAATG 6676
DB 2508 TTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGT 2567
QY 6677 CGGAGTCGAGTGGCTGCAAAAAGAGCTGGAATTTGGAGCACAGTGGCAAGTTACCAAG 6736
DB 2568 TAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAACTG 2627
QY 6737 ATGAATCAAAACAGAGAGCAAGTGAATAATTTCTAAAATTT--ATGACTCGAAACAACT 6793
DB 2628 TTGAAGAAATTTAGCTCCAACTGTTGAAGAAAGTGTAGCTCCAACTGTTGAAGAAAGT 2687
QY 6794 TAGATGTAAATGACAGAAATGAGATTAATCAGGTACTGGAATCGGTTACGCCGAGCTG 6853
DB 2688 TTGCTGAAAACGTTGCAACAAATTTATCAGCAATCTTTTAAGTAAATTTATAGGTGGA 2747
QY 6854 GAATCTTGACAGCCGAGATCTGGAGTGGTTCTGTCAATTAATATGCAAAATTAAGTGA 6913
DB 2748 TCGAAACTGAGAAATTAAGACAGTATATTAATATGAGATGAGAGAAAGTAAAGAAATG 2807
QY 6914 AAACAGATATCATCTACTACTTTTACACTCTTCTACTAGATGTAATGTAAGAAAGTCTTA 6973
DB 2808 TAGTCACCAATCTAGAAAACGTAGAGAAAGTGTAGCTGAAAGTGTAACTACTTTTA 2867
QY 6974 ATTAATTTTCGAATCTCTTGACAGCCGCTGGAGAGCCGAGCTCTTGACAGATTTACCG 7033
DB 2868 GTTAATATATTAGAGAGATACAAAGAAATATCTAATGATATCTAATAGAGAAAT 2927
QY 7034 GAGTGGTTCTGTAACTATTAATAGTTCGTGATAGCTGAGTTCACATTAATCTG 7093
DB 2928 TGAAGAACTCCAGCAAAATGTTATGAGTGGCTTTAGAAAATGCCAACTGAGAGG 2987
QY 7094 ATTTGACTCCGTAACGAGA--AAAAGTAATGTAAACGCAAAAGAGAAAATATATTA 7150
DB 2988 AAAAGAAAGATTAATGATGTATTTGAAGAAAGTAAAGAGAGTGGTACCACTTTA 3047
QY 7151 AGCAAAACAGCAAAATGACAGAAATCGAGAGACAGCAATCGAGCCAAATGTTGTGTA 7210
DB 3048 TAGAAAATCTGTGAAACGCGAGAAAGAAAGCGCAATTCAGTAATATTTGAAA 3107
QY 7211 ATTAATTTTGGACAGCTGTAGAAAGATGAAAAATTTGGAAGAAAGAAAGAAAGTT 7270
DB 3108 ATTTGAGAGAAATGCAAGTGAAGTATGAAAATGTTGCAAGAAATTTAGAGAAATTA 3167
QY 7271 TAAAAACTTTAGACGAAAGTTAAACAAGAAACAAGATTAAGAAAGTAAATGATGCTACGAAA 7330
DB 3168 ACGAAAGCTATTTAATACGTATTAAGATAAGTGAAGAAACAAGTAAATTTACCGAG 3227
QY 7331 AAATCTTACATCAGCAGGATTTTCTACAGAAAGTACTTGTAAAGCGGATGAAGAG 7390
DB 3228 AAAAGTTAGAAAACAATGAAATGAAATGCAATTTTGTAGAAAATATTTGTAATGTA 3287
QY 7391 AATACCAAGGAGAAAGAAATTAAGCATTTGGAAGCTTGTATATTAATGAAAAATG 7450
DB 3288 AAGGAATACAGAAATTTATTAACAGGTATG-----TTGGAAGTATAGAAACACAGTA 3341
QY 7451 TAGATATTACACAGAGACAAGAAATATATCACTTCTAGTGGTGTGGAACTGACG 7510
DB 3342 TAGATATCAATCAGAAAGAAAGGTTGATTTGAAATGAAAAATGTTGATGATTTTGA 3401
QY 7511 GTCTTGCTTCCGATCAGGAACAGTGGCAGTTACAAATATTTAAAGAAATTCGCGAGTTA 7570

DB 3402 ATTAATATGAAAAATATGAAAAAGGTTTATTAATTAATTAATTAATTAATTTCAAGTACTG 3461
QY 7571 CTGTTGAAAATTTCTTTGTTGAAGACGCTGAAAAAGTAAATGTTAGATCGGATAT----- 7625
DB 3462 AAGGTTCACAGAACTGTAATCTGAACTGTGAACATGTGAACAAAATGATATGTGATGTTGATG 3521
QY 7626 ----TACAGAAATGTTCTTTTAAACAGCATATCAAGGCTCTGTAGAGACATTTGGAAATG 7681
DB 3522 TTCTGCTATGAAAGATCAATTTTGAAGATTAATTAATTAAGAGCAGAGGTTGAAGAAA 3581
QY 7682 GAGCTGCTATGCAAGATTAATTTCTAATGGAAGTCAATATCATGTTAATAATTTCTA 7741
DB 3582 TGTTTTAAATTTGGAAGATGATTTAAAGTGAAGTGAATGATTAATCTGTAGAGAAA 3641
QY 7742 AGCTATTAGAAAAAATTTGATGTTATGTTAAAGATTAATCGGAATTGAGAGCGAAG 7801
DB 3642 TTAAGATGAACCGGTTCAAAAAGAGTAGAAAAGAACTGTTAGTATTTATGGAAGAA 3701
QY 7802 CAAGAAGTTAACGTTAGAGCCGTTAGCTGCCGAGCCATTATCTCAAAAGCAAGAAATG 7861
DB 3702 TGAAGAAATATTTGATATGATATTAAGAGAAAGAAAGAAAGTTTAACAAGCAAGATGA 3761
QY 7862 AAATGAATTCAGAGTTGAAATTTGAGAAAGTATTTTCAATGAAGAAATGAGTAACTA 7921
DB 3762 TAGATGAGTGAAGATCCATGAAATATCTCAGATTTCTAAGAGAAAGAACTGAATCTA 3821
QY 7922 GCCCTTTAAAGGAATTGGAAGAGAAATCAATGTCAAAAGTGGAAAAAGAAACAGAGTA 7981
DB 3822 TTAAGATTAAGAAAGAAAGATGTTTCACTAGTGTGGAAGAAAGTTCAAGACATGATATG 3881
QY 7982 CTGCTGAATCTCAAGAGACT-----TCTGTAGAGCAGTGAAGCGGCGAG 8027
DB 3882 ATGAAGAGTGTGAGAAAGTTTAAATGAAATATGAAAGAGAGATTAATGAAGATG 3941
QY 8028 AATTAATTTCCAGAGCAAAAGTCCGGAAGCTCTTATTTGAAGTATGATACAAATCCGG 8087
DB 3942 CTGTGAAATTAAGACACTTCTAGCAAACTTATGGAAGAACTCAAGAGTTAAATGAG 4001
QY 8088 AAGAAGTATTTTTCATGAGATAATGTAATGTAATGGAAGCAACATTAATGAAGTAAG 8147
DB 4002 TGAAGCAGATTTAATTAAGATTTGAAAATTTAAGAAATTTAGAAAAGC--ATTAATC 4059
QY 8148 AGCAGTTTCTAAGCAGTACAGGTTCTGTATTTGGAGAGTGGAGTCAACCAAGCAGA 8207
DB 4060 AAGAGATTTCTAAGAAATTAATGATGCAAAAGATGATCATTTAGAAAAGTTATGAAGA 4119
QY 8208 AGCTACTGCTCAGGTAAACTATGTAGAAAGTTGAGAGAAAGAAATTTGTTCAAGACAA 8267
DB 4120 GGAACATGATATTAACGACGAGCTGGATGAAGTGTAGAAATTAAGAAAGTGTGAAGAGA 4179
QY 8268 TCGATGGAATGCAATTTCTAAGTGAAGGTTTGAATGAAGATTAAGTAACTGCTAAATC 8327
DB 4180 CAGATGCAAAAAGTATGTTGATTTAAGAGATCTTGAAGAAAGATATTTAAG----- 4232
QY 8328 TTCTGTAGTATCAGAAATGAGAGAGAAATTCGCGAGCAGAGTGAATACTTCAAGC 8387
DB 4233 --AGTAAAGAAATCAAAAGAACTGAAAGGAATTTTAAGATTTAAGAAATTA 4290
QY 8388 ACAAGTAAATCTGAATCCGATGCTTTACGAAGCAGAAATTAAGAAATTAATGATTA 8447
DB 4291 AACTATTGAAACAATATTTTGAAGAGAAAAAGAAATTAAGAAAGATCAATTTTGA 4350
QY 8448 CACAAAAAATATATTTTCAAGAGTCAATGCTGTTGATTAATGATACAAAGAAATGAAGC 8507
DB 4351 ATTGAAGAAAGCTGAAGAAATTAAGATCTTGAAGCAATATATTTAAAGAAAGTATC 4410
QY 8508 GAATATGAATCTTTAGCGGTAGCCGCTGTGATGCAACAAGAAACAAGCAAGATTTAC 8567
DB 4411 TTCAATTAAGATTTGAAGAAAGAAAGAAATTTGAAGAAAGTACAGAAATTAAGAAAGAGT 4470
QY 8568 GAGATCAAAACAGTTAACTTTACACTGTAAATGAGAGAAACGATCTCAACTTCGTC 8627

Db 4471 AGACATATATATATGATGCGCATATATAAGGTTGGAGAAAGATGATTAGAAGA 4530
Qy 8628 AAAAGCTTGGCTAAAAATGAAAATTATGAAATGTAAAGAACTGAGAGACCTTACT 8687
Db 4531 AGTAGATGA-ITTTAAAGAAAGATATATGACATGTTAAAGGAGATATGAG--ATTAGG 4587
Qy 8668 CGGAGCGGAAACAGAGCCGTTGAAAATTATCAAAAGAGTACTACAGAGGACCTTGGC 8747
Db 4588 GATATGATTAAGAAAGTTAGAGATGTAACCAACAACTTGGAGAAAGATTGAATC 4647
Qy 8748 AGGAATTTGGAAATTTGAGATTAATTAAGAAACGATTCAGAGATTAATACGATTGTAG 8807
Db 4648 CTTAAAGATGTTTATCTAGTCATTTAGGATGATGAAAGAACAAATGAAACAAAGAA 4707
Qy 8808 AGTCAACGAGACGAAACCAAGAGCTTGTGCGAAAGAAATGTTCTGTGAAATA 8867
Db 4708 AAAAGCTCAAGACCTTAAGTTGGAAGATATTTAAAGAAAGAGTTAAAGAAAGAAC 4767
Qy 8868 TACAAATTTGAGGGGAAACAAATTCATCCATTGAAGATTAAGCCAGATTTGTTGAAACGG 8927
Db 4768 AAAAGAAATTAACCAAAAGAAAGTAAAGTTGATTAAGATTAAGAAACCAAGAA 4827
Qy 8928 AAGTGAATATGATGCTTTGATGAACTTGTGATGATCTACAAAGAAAGTGTGG 8987
Db 4828 TGAATATGATGAAGTTGAAATGAAAGATGAAGTATGAAGAAAGATGTGAAGAAATAT 4887
Qy 8988 CTATGCTGATTTGATGATGAAATGTTGATGATTAATTAATGATTAAGAAAGATGTA 9047
Db 4888 AGAAGAGATTAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 4947
Qy 9048 AGCAGAAATCGAGACATGCTATTTGATTAAGAACTACTGAGAAACAGAAATTCAGCAAT 9107
Db 4948 AGGTGAAGCAAGATGAAGTATTAAGTCAAAAGAAAGAAAGCAAGTTGAAAA 5007
Qy 9108 TACAAGAGCAAGTAAATATTTCTGAAAGAGAGCGCTGACGCTGACATGATATC 9167
Db 5008 GGTAAAGGCAAAAGAAATTAAGAAAGAAAGTTGAAAGAGTGTAGTGTCTTAA 5067
Qy 9168 GAATGACATTTCCATGAGATGATTAAGAAATTTGGCAAGAGATGATGCTTC 9227
Db 5068 AAAACGATGACAGAGTATGAATATGTTCAAAATTTGATTAAGAAAGTGTATGAAGA 5127
Qy 9228 TCAATTAATTAACCAAAATTTCAAAATTAATTAATTAATTAATTAATTAATTAAT 9274
Db 5128 AGTATCTTAAGCTTTAGAAATCAAAATTAATTAATTAATTAATTAATTAATTAATTA 5174

RESULT 11
AAT05868 standard; DNA; 3399 BP.
ID AAT05868;
XX
AC AAT05868;
XX
DT 27-AUG-2003 (revised)
DT 14-AUG-1996 (first entry)
XX
DB Chicken leucocytozoan DNA encoding immunogenic protein for vaccines.
XX
KW Chicken leucocytozoan; immunogen; recombinant vaccine; protection;
XX immunisation; vaccination; ss.
XX
OS Leucocytozoan sp.
XX
XX
XX Key Location/Qualifiers
XX PH 1..3399
XX FT CDS
XX FT misc_feature
XX FT /tag= a
XX FT /tag= b
XX FT /note= "fragment referred to in the claims, for use as
XX FT insert in a recombinant vaccine against chicken
XX FT leucocytozoan disease"
XX
XX PN JP07284392-A.

XX 31-OCT-1995.
PD 19-APR-1994; 94JP-00080643.
XX 19-APR-1994; 94JP-00080643.
XX 19-APR-1994; 94JP-00080643.
XX (DOBU-) DOBUTSUO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.
XX (KITA) KITASATO KENKUSHO SH.
XX WPI: 1996-006311/01.
XX P-PSDB; AAR97866.
XX
XX Chicken leucocytozoan immunogenic protein - used in a recombinant vaccine
XX against chicken leucocytozoan disease.
XX
XX Claim 6; Page 6-9; 35pp; Japanese.
XX
XX AAT05868 encodes a chicken leucocytozoan immunogenic protein, this DNA or
XX a fragment of it can be used in a recombinant vaccine to immunise against
XX chicken leucocytozoan disease. The DNA is used in a vector and
XX CC operatively linked to an expression regulatory sequence as in standard
XX CC practice. (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 3399 BP; 1577 A; 508 C; 798 G; 516 T; 0 U; 0 Other;
SQ
Query Match 1.1%; Score 104.2; DB 2; Length 3399;
Best Local Similarity 43.3%; Pred. No. 4.8e-10;
Matches 748; Conservative 0; Mismatches 963; Indels 16; Gaps 5;
Qy 7110 AGAAAGATTAATGATGACGCAAAAGAGAAATTAATTAAGCAACGACGCAATGC 7169
Db 1461 AAACACACATCTGAAACAGAAATTAACATCTATGAAACGATTAACAAATGA 1520
Qy 7170 AGCAATCGAGAGAGAGCAATCGCAATGCTTTGTAATTAATTTGCAACGCTT 7229
Db 1521 TGAAGAAAGCAACACATTAATGAAGAAATTAATTAATTAATTAATTAATTAATTA 1580
Qy 7230 AGAATGATGAATTAATTTCTGAAAGAAAGAAAGCAAGTTTAAATCTTACAGCAAT 7289
Db 1581 AGTAAATGATGAAGATCGATGAAGCAAGATTAATTAATTAATTAATTAATTAATTA 1640
Qy 7290 TAAACAAGAAACAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7349
Db 1641 ACAGAGAGAGCTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1700
Qy 7350 TATTTCTACAGAGATCTTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7409
Db 1701 AATAGAAAGAAAGAGATGAAAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 1760
Qy 7410 TAAAGCATTTGAGAGATCTTCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 7469
Db 1761 TGAAGAAATTAAGAAAGAAAGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1820
Qy 7470 CAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7529
Db 1821 AACACATGAAGAAAGAAAGAAAGAGATTAATTAATTAATTAATTAATTAATTAATTA 1880
Qy 7530 AACAGTGCAGTT--ACAATATTAAGAAATTTCCGAGTTACTGTTGAAATTTCTTT 7587
Db 1881 AGAAGTAAACATGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1940
Qy 7588 GTGAAGACAGCTGAAGAAAGTAAATGTTAGTGGATTAATTAATTAATTAATTAATTA 7647
Db 1941 AAAAGAAAGATTAACATGAAGAAATTAAGAAAGAAAGAGATTAATTAATTAATTAAT 2000
Qy 7648 GCATATCAAGGCTCTGATGAGATGGAATTAATTAATTAATTAATTAATTAATTAATTA 7707
Db 2001 GGAAGAAAGAAAGAAAGTAAACATGAAGAAATTAATTAATTAATTAATTAATTAATTA 2060
Qy 7708 AATGAAGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7767
Db 2061 ACATGAAGAAAGAAAGAAAGTAAACATGAAG--AATTAAGAAAGAAAGAGATGA 2118


```
QY 7232 AAGTAGAAAAATTTGGAAGAAAAGGACAGAGTTTAAAACTTTAGACGAGTTA 7291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 TCTTACTGAAAAATPAAAGAACAAAAATGAAAGATCCAAATGAAATAGAGTAGTAA 349
QY 7292 ACAAGAAACAAGATPAAAAAGTAAATGATGCTACGAAAAAAATCTTCAATCAGCAGSTA 7351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 ATGATGAGAAAGAGTAAAAACAGATTTGATCTGAAAAAAAATGAGAGATGAAAAATA 409
QY 7352 TTTCTACAGAGATCTTCTGTAAAAAGCGGATAGAGAGATCTCAGGAGAGAGATTA 7411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 AATCGAAAACTGAAATAGGTGAGAAATTAACGTAAAAAGTAAATGAAAAAGTACCTGAAG 469
QY 7412 AAGCATTTGGAAGCTTCTGATTTATTTGAAAAAATGATGATTTTACAACAGAGACA 7471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 AAGTAGCTGAAGATTAAGTAAAAAGTAGAGAAAGTAGAGAGTGAAGATTAAGTTGAAA 529
QY 7472 AGAATTAATATCACTTCTCTGATGTTGGGAACTGCAGATCTTGCTTCCGCATCAGGAA 7531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 AAGTAGATGAAAAAGTACGTGAAGAGTAGATCAAAAAAGTAGATGAAAGATTAAGTGAAG 589
QY 7532 CAGTGGCAGTTACAAATATTAAGAAAAATCCGAGTTACTGTTGAAAAATCTTTTGCA 7591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 590 AATTAAATGAAAAAGTAGATGAAGAGTAAGTGAAGATTAATGAAAAAGTAGATGAAG 649
QY 7592 AAGCAGCTGAAAAATTAATGTTAGATCGGATATTACGAAAAATGCTTTAACAGCAT 7651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 650 AAGTGTCTGAAGAAATTAATGAAAAAGTAGAGAAAGTGTGCGAAGAAATTAATGAAA 709
QY 7652 ATCAAGGTCTG---TAGAGACATTGGGAATAGAGTGCCTATGTCAGAAATTAATTTCTA 7708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 710 AGGTAGCTGATGAATTAATGAAAAAGTAGAGAAAGTGTGCTGAAGAAATTAATTTGAAA 769
QY 7709 ATGGAAGATCAAAATATCAGATTTAAAAATTTCTAGCTATTAGAAAAAAATATGATGTTA 7768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 770 AGTAGCTGAGAAATTAAGTTGAAAAAGTAGAGTGAAGAAATTAAGTTGAAAAAGTAGAGTGAAG 829
QY 7769 TTGTAAAGATTAATCGGATTTGAGACGGGAGCAAAAGATTAACCTAGAGCGGTAG 7828
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 830 AAGTAGCTGAAGAAATTAAGTTGAAAAAGTAGAGTGAAGAAAGTGTGAGAAAGTAGATCAA 889
QY 7829 CTGCGGAGCCATTTCTCAAAAGCAAAAGATGAATGAATTCAGAGGTGAAATTTGAGA 7888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 890 AAGTAGATGAAGAAATTAAGTGAAGAAATTAATTTGAAAAAGTAGAGTGAAGATTAAGTGAAG 949
QY 7889 AGAGTATTTTCAATGAAGAAAAATAGATTAAGTCCCTTTCTAAAGAAATTTGAGAGAAA 7948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 950 AATTATTTGAAAAAGTAGATGAAGAAAGTGTGAGAAATTAATTTGAAAAAGTAGATGAAG 1009
QY 7949 TCAATGTCAAAGTGAAAAAGAAACAGAGTACGTGATCTCAAGAGCTTCTGTAG 8008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1010 AAGTGTCTGAAGAAATTAATTTGAAAAAGTAGATGAATTAAGTTGAAAAAGTAGATGAAG 1069
QY 8009 GAGCGATGACAGGCGAGGAATTAATTTCCGAGACAAAGATCCGGAAGCTCTTAATTTGA 8068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1070 AATTAGTTGAAAAAGTAGATGAAGAAAGTGTGAGAAATTAAGTTGAAAAAGTAGATGAAG 1129
QY 8069 AAGTTAGTACAAATTCGAGAGAGATTTTTCATGACAGATTAATGAAATATGAGAACAA 8128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1130 AAGTAGCTGAAGAAATTAAGTTGAAAAAGTAGATGAAGAAATTAAGTTGAAAAAGTAGATGAAG 1189
QY 8129 CACATTAATGAAGATTAACAGCAGTTTCTAAAGCAGTAAACAGTTCTGTATTGGAGAGAG 8188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1190 AAGTAGCTGAAGAGATTTGAAAGAGTGAAGAAAGTGTGAGAAAGTAGATGAAGAG 1249
QY 8190 AAGTAGCTGAAGAGATTTGAAAGAGTGAAGAAAGTGTGAGAAAGTAGATGAAGAG 1249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8189 TTGAGATCAACAAGCAGAGAGTACTGCTGAGGTAAATCTATGAGAGAGTTTGGAGAG 8248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1250 TAGCTGAGAGAGTACTGAGAAAGTAGTGAAGAGATGACTGAAGAAATTAAGTTGAAAAAG 1309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8249 GAATTTGTTTCAAGAAC---AAATCGATTTGAAATGCAATTTCTAAAGTAGAGAGTTTGATG 8305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1310 TAGATGAAGAGATGCTGAAGAAAGTAGTGAAGAAAGTAGTGAAGAAAGTAGTGAAGAG 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8306 AAGATTAAGTAACCTGTAATCTTCTGTGTATCAGAGAAATGAGAGAGAAATTTGCCGAG 8365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1370 TAGTTGAAGAGATGATGAAGAGATGAGTGAAGAAAGTAGTGAAGAGAGAGTGAAGAAAG 1429
QY 8366 CAGAGTGAATATCTTCTACAGCACAAAGTAACTAGTAATCCGTAGTCTTTAGGAAAGC 8425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1430 TACTTGAAGAGATTAATTAAGAGATGATGAGAAAGTAGAGCGAAGAGTACTGAAAAAG 1489
QY 8426 AAGATTGAAAAATTAATGATTAACAAAAAATATATTTGAGAG 8470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1490 TAGTTGAAGAGCAGAGTGAAGAAAGTAAACAAAAATGATTTAAATG 1534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
AAA70099
ID AAA70099 standard; DNA; 3579 BP.
XX
AC AAA70099;
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:232.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KM antimalarial; malaria; protozoacide; infection; insecticide; de.
XX
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US026796.
XX
PR 05-NOV-1998; 98US-0107131P.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
XX Plasmodium falciparum, useful as anti-malarial vaccines and in the
XX diagnosis of P.falciparum infection.
XX
PS Disclosure; Page 457-458; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II). (I) and
CC (II) are useful for the development of vaccines against P. falciparum
CC infection. (I) and polyclonal antisera or a monoclonal antibody raised to
CC immunogens comprising the sequences of (I), are useful in the detection
CC of infection with P. falciparum. Furthermore, (I) (especially when they
CC are rifins or secreted or membrane proteins) can aid the identification
CC of drugs to treat or prevent P. falciparum infection, or they can be used
CC to identify drug resistance in P. falciparum. Sequencing of the
CC Plasmodium chromosome 2 and the subsequent identification of proteins
CC encoded by it will help to expand our understanding of parasite biology,
CC a process hampered by the complexity of the parasitic lifecycle, and
CC provide new targets for vaccine and drug development. Parasite resistance
CC to drugs and mosquito resistance to insecticides have led to a resurgence
CC of malaria in many parts of the world, and there is a pressing need for
CC vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
CC represent nucleotide and protein sequences given in the present
CC invention, but which are not specifically mentioned within the
XX specification
XX
Sequence 3579 BP; 1904 A; 398 C; 552 G; 725 T; 0 U; 0 Other;
```

Query Match	0.9%;	Score 87.4;	DB 3;	Length 3579;
Best Local S.Milarity	42.7%;	Pred. No. 9e-07;		
Matches	616;	Conservative	0;	Mismatches 821; Indels 6; Gaps 3;
Oy	7110	AGAAAAAGTAAATGTAACGGCAAAAGAGAAAAAATTAAGCAAAACAGACGCAAAATGC	7169	
Db	1362	AGAAATCAGAAAAAGATTAAGAAAGATTCAGAAAAAGACAAAGAAAAAATCGAAGAAAGATGA	1421	
Oy	7170	AGGAATCGAGAGAGAGCAATCCGAGCCAAATCTTGGTAAATATTTTGGAAACGCTGT	7229	
Db	1422	AGAAAAAATCTGAAGCGAAAAAGAACGAAAGTATACAAAAAGAAACAGATGTGATGA	1481	
Oy	7230	AGAAAGTAAAAAATTTCTGAAGAAAAAGAAACAGAAATTTTAAAAAATTTAGACGAAGT	7289	
Db	1482	AAAAAAGAAAAAGGAGAAATATGAGAGGGAACGATGATGAGAAAGCAAAAGAAAAAGAA	1541	
Oy	7290	TAAACAAAGACAGATTAATAAAAGTAATGATGCTACGAAAAAATCTTACATTCAGCAGG	7349	
Db	1542	AGAAAGACGACGAAACAAAGTAAAGAAAAAGAAAAACAGAAAAAGACGAAAGAGGAAAC	1601	
Oy	7350	TATTTCTACAGAAATCTTCTGTAAAAACCGATAGAGAGATCTCAGGAGAAAGAAAT	7409	
Db	1602	AGATTATAGAGAAATACAGATGATTCAGACAAAGATGAAAGAAACAAAGTAAAGAAAA	1661	
Oy	7410	TAAAGCCATTGGAAGACTTCTGATATTATGAAAAAATGTAGATTTACACAGAGGA	7469	
Db	1662	GAAAAACAGAAAGACGAAAGAAAGAACTGAAGAAACGAAAAAGAAACAGAAAGTGAAGAA	1721	
Oy	7470	CAGAAATATATCACTTCTACTGTGTGGAACTGCAAGCTCTTCCGATCAGG	7529	
Db	1722	AAAGAAAAACAGAAAAAGACGAAAGAGGAAACAGATTATGAAAGAAATACAGATGATTCAGA	1781	
Oy	7530	AAAGTGGCAGTTACAAATATTTAA--AGAAATTCGGAAGTACTGTTGAAAAATTTCTT	7586	
Db	1782	CAAAGATGTAGAAACAGAAAGTGAAGAAACAGACCGCAGAAAGCAAAAGAAAGAAAGA	1841	
Oy	7587	TGTGAAACGACGTGAAAAAGTAAATGTTAGATCGATATTACAGGAAATGTGCTTTAAC	7646	
Db	1842	GGGAACAGATGATGAAAGAAAGACAAAGTAAAGAAACACACCTTAGCGAACAAAGAAAGA	1901	
Oy	7647	AGCATATCAAGTCTGTAGAGCACTTGGAATAGAGACCTGCTATGAGAAATTTAAATTC	7706	
Db	1902	CGGAGAAAGAGTAAAGAAAGACGAAAGAAAAAAGACAAAGAAAGATGACAAAGAAAGATGA	1961	
Oy	7707	TAAATGAGATCAATATCAGTATTTAAATTTCTAAGCTATTAGAAAAAATATTGATGT	7766	
Db	1962	CAAAGAAAAAGACAAAGAAAGATGACAAAGAAAAATACAAAGAAAGATGACAAAGAAAGATGA	2021	
Oy	7767	TATTTGTAAGAATTAATCGAAATTTGAGACGGAACCAAAAGATTTAACCGTAGAGCGGT	7826	
Db	2022	CAAAGAAATGTGACAAAGAAAAAGACAAAGAAAGATTAACAAAGAAAAAGCAAAAGAAATTA	2081	
Oy	7827	AGCTGCCGAGCCATTATCTCAAAAAGCAAAAGATTAATGAATTCAGAGTTGAAATTTGA	7886	
Db	2082	CAAAGAAAAAGACAAAGAAAGATGACAAAGAAAAAAGACAAAGAAAGATGACAAAGAAAAAGA	2141	
Oy	7887	GAAAGTATTTTCAATGAAAGAAATAGATTAAGCCCTTTAAAGAAATTTGGAAGGA	7946	
Db	2142	CAAAGAAATTAACAAAGAAAAAGACAAAGAAAGATTAACAAAGAAAAAGCAAAAGAAAGATGA	2201	
Oy	7947	AATCAATGTCAAAAGTGAAGAAAAAGAAACAGAGTGACTGCTGAATCTCAAG--GAGCTTCT	8004	
Db	2202	CAAAGAAAAAAGCTGTAATTAACATGTAAAGAAATTAAGAAAGAAATGAAAGATGATGACTA	2261	
Oy	8005	GTAGAGCAGTAGACAGGGCAGAGAAATTAATTTCCGAGCAAAAGATGCCGGAAGCTTTAT	8064	
Db	2262	TGATGAAAGTTTGAAGAAAGAAATTAATTCACATTAATATAGACATTTGGCAACACACA	2321	
Oy	8065	TTGAAAGTTAGTACAAATCCGAAAGAAATTTTTCATGCGAGTAATGTGAATTTGGA	8124	
Db	2322	ATATTTCCATCTACTAATCAATCTTGAACCAAGAAATTAATCAACTATATTAATAA	2381	

Oy	8125	GCAACACATATAAATGAAAGATTAACAGCAGTTTCTAAAGCAGTAAGGTTCTGTATTGGGA	8184	
Db	2382	ACAACAGAAAAATTAATCAACCCACATCATATTATTAACCAACAGCAAAATCATTAATCCACA	2441	
Oy	8185	GGAAGTTGAGTCCACCAAGGCAAGAGCTACTGCTCAGGTAATAACTATGTGTGAAGTTGAG	8244	
Db	2442	TCATATTTTACAGAGCAAGAAAAACATCACCCACAGGTATACCAAGGAAGCAACCATTA	2501	
Oy	8245	GAAGAAATTTGTTCAAGAACAATCGATGAAATGCAATTTCTAAAGTAAAGGTT--TGGA	8303	
Db	2502	TAAATATGTTCTTATATATCTTAAAAAAGGGGCTTGAACCCAAATCTCATACATGTAA	2561	
Oy	8304	TGAAGATTAAGTAACTGCTAAATCTTCTGTAGATTCAGAAATGAGAGGAAATTTGCCGG	8363	
Db	2562	AGAAAGTCAACTATATTAATAACAGGGGTTGTAAAGGGAACAAGAACACATGTGATGA	2621	
Oy	8364	AGCAGAGTGAATCTTCTACAGCAACAAGTAATCTGAATTCGGTATGCTTTACGAA	8423	
Db	2622	TATGCATACAAATCAAAAGAACATAGAAATTTTAAAAATGCAACCGATGTAAACCAACC	2681	
Oy	8424	GCAAGATTATGAATAATATGATTAACAAAAAATATTTTCAAGAGTCAATGCTCTTG	8483	
Db	2682	AGCAAGTCATATATATTAATTAATTCATCGAAAAACAAATTAACATGTATTAATAATC	2741	
Oy	8484	TTTAATGATACAAAGATGAGGCAATATAGAACTTTAGCGGTAGCCGGTGTGCATGC	8543	
Db	2742	TCCTGAAAAACAAATTTGAACATGTATATTAATTAATGCTCTCTAAAAACAAATTTGAACATGT	2801	
Oy	8544	ACA 8546		
Db	2802	ATA 2804		
RESULT 14				
AA087587	ID AA087587 standard; DNA; 1686 BP.			
XX	XX	XX	XX	XX
AC	AA087587;			
XX	XX	XX	XX	XX
DT	27-AUG-2003 (revised)			
DT	19-DEC-1995 (first entry)			
XX	XX	XX	XX	XX
DE	DNA encoding Leucocytozoan protozoa structural protein epitope.			
XX	XX	XX	XX	XX
KW	leucocytozoan protozoa; structural protein; epitope; vaccine; fowl;			
KW	leucocytozoanosis; treatment; ss.			
OS	Leucocytozoon.			
XX	XX	XX	XX	XX
PN	JP0708995-A.			
XX	XX	XX	XX	XX
PD	04-APR-1995.			
XX	XX	XX	XX	XX
PF	10-SEP-1993; 93JP-00226078.			
XX	XX	XX	XX	XX
PR	10-SEP-1993; 93JP-00226078.			
XX	XX	XX	XX	XX
PA	(DOBU-) DOBUTSURO SEIBUTSUOKUTEKI SEIZAI KYOKAI.			
XX	XX	XX	XX	XX
PA	(NISS-) NISSEIKEN KK.			
XX	XX	XX	XX	XX
DR	WPI; 1995-167252/22.			
XX	XX	XX	XX	XX
DR	P-PSDB; AAR70491.			
XX	XX	XX	XX	XX
PT	Immunizing polypeptide against Leucocytozoan protozoa - useful in			
XX	XX	XX	XX	XX
PT	production of vaccines for treatment of Leucocytozoanosis in fowl.			
XX	XX	XX	XX	XX
PS	Claim 1; Page 12-14; 20p; Japanese.			
XX	XX	XX	XX	XX
CC	AA087587-89 encode polypeptides having a whole or partial epitope of a			
CC	structural protein of Leucocytozoan protozoa (see AAR70491-93). The			
CC	polypeptides and DNA encoding them are useful in the production of			
CC	vaccines for the treatment of Leucocytozoanosis of fowl. (Updated on 27-			
CC	AUG-2003 to correct OS field.)			

XX Sequence 1686 BP; 915 A; 129 C; 382 G; 260 T; 0 U; 0 Other;
SQ

Query Match 0.84; Score 81.8; DB 2; Length 1686;

Best Local Similarity 42.4%; Pred. No. 8.8e-06;

Matches 513; Conservative 0; Mismatches 692; Indels 6; Gaps 1;

OY	7112	AAAAAGTAAATGTAACGCCAAATAAGAAAAAATATTATTAAGCAACAGCGCAAAATCCAG	7171
Db	302	AAGATGAAGAAGAACAGAAAGAAAGAAACAGATGAAGAGAAAGAAAGAAAGATG	361
OY	7172	GAATCGGAGGAGCAGCAATCCGAGCCCAATGTCCTTGTAATTAATTTTGAAACAGCTGAG	7231
Db	362	AAGAAAGAAAGAAAGAAAGAAAGAAACAAATATGAAGAAACAAATATGAATGAAACAAATG	421
OY	7232	AAGATAGAAAAAATTTCTGAAGAAAAGGAACAGAAATTTTAAATCTTTAGACGAATTA	7291
Db	422	AAGATGAACAAAATGAAAGMAGMAGMAGMAGMAGMAGMAGMAGMAGMAGMAGMAGMAG	481
OY	7292	ACAAAGAACAAAGATTAATAAAGTAAATGATGCTACGAAAAAAATCTTCAATCAGCAGTA	7351
Db	482	AAGAAAGAACAAAGATGAAGAAACAGATGGAAGAAAGAAAGAAAGAAAGAAAGAAAG	541
OY	7352	TTTCTACAGAAAGATATCTTCTGTAAAGCGGATAGAGAGATATCTCAGGAGAAAGAAATTA	7411
Db	542	AAGATGAAGAACAAAGAAAGATATATGCTGAAAAAAGAAATGAAAGATGAAGAAAAAAG	601
OY	7412	AAGCATTGTGAAGACTTCTGATATTATTGAAAAAAATGTAGATATTACACAGAGACA	7471
Db	602	AAAAAGAAAGAAACMAGMAGATGAAAAATATATGTGAAAAAGAAAAAAGTAAGAG	661
OY	7472	AGAAATATATCACTTCTACTGATGTTGGGAACTGCGAGTCTTGCTCCGATCAGGAA	7531
Db	662	TATATGTGAAGAAAGAAACAAAGACATTTTAAAGATTAAGAAAGAAAGATTTAGAGAA	721
OY	7532	CAGTGGCAGTTACAAATTTTAAAGAAATCCGAGATTACTGTGAAAAATCTTTTGCA	7591
Db	722	AAGTGAAGAGTGAACATTAAGATGATTAAGAAAGAACAAAAATAAAGCATAGAGTG	781
OY	7592	AAGCAGCTGAAAAAGTAAATGTTAGATCGGATATTACAGAAATGTTGCTTTACAGCAT	7651
Db	782	AAGAAAGAACGAAATAGAAATTAAGAAAGAAATCAGAAAGAAATAGATAGAAAAATCAG	841
OY	7652	ATCAAGTCTCTGAGAGACATTTGGGAAATAGAGCTGCTATGCAAAATTAATTCTAATG	7711
Db	842	AAGTAAAAAGAACAAATATGAGCAAGTACAAAGAAAAACAAATGACAAAAAGATATATAC	901
OY	7712	GAAATCAAAATATCAATTTTAAATTTCTAAGCTTTAGGAAAAAATTTGATGTTATG	7771
Db	902	AAGATCAAAATTTGGTCATCAATATTATAGATATATATTAATTAAGAACACAGAAACAGAA	961
OY	7772	TAAAGATTAATCCGAATTGAGAGCGGAAGCAAAAGATTAACGTAGAGACGGTAGCTG	7831
Db	962	AAACTTAATCAGAAATTTATACATTAAGTAACTTAATGAATCAAAATCCGAATTAATT	1021
OY	7832	CCGAGCCATTATCTCAAAAGCAAAGATGAATTCAGAGTTGAATTGAGAA	7891
Db	1022	CAACTACAAATATTAATTTGAAAGTGAATAAATTTCAATCTTGAAAAAACACAGATTTAA	1081
OY	7892	GTATTTTCAATGAAGAAAAATAGATTAATCAGCCCTTCAAAGGAATTGGAAGAGAAATCA	7951
Db	1082	ATGAAGCTACTAGAGCGCGGAATTCGGACACAGTATGGAAGAAATTTTGAAGAAACAC	1141
OY	7952	ATGTCAAAAGTGAAGAAAAAGAAACAGAGTGAATCTTCAGAGACTTCTGAGAG	8011
Db	1142	CAGTATGTTGAAGAAAGAAATGATGAAGAAACACTGTGATGAAAGAAATGATAGAAG	1201
OY	8012	CA-----GTACAGGGCAGAAATTAATTTCCGAACAAAGATCCCGGAAGCTTTAT	8066
Db	1202	AACACACTGTATGTTGAAGAAAGAAATGATGAAGAAACACCAATATGGAAGAAAGTTCG	1261
OY	8066	TGAAGTTAGTCAAAATCCGAAAGAAAGTATTTTTCATGACAGATATGTGATATGAGAG	8125

[illegible]

RESULT 15

ADP85917

ID ADP85917 standard; DNA; 1300 BP.

AC ADP85917;

DT 26-AUG-2004 (first entry)

Synthetic construct #1.

KW Nanopore data analysis ; polymer ; polynucleotide

single-molecule resolution; ds.

OS Synthetic.

PN US2004110205-A1

PD 10-JUN-2004

23-SEP-2003; 2003US-00668749.

23-SEP-2002; 2002US-0412959P.

PA (WANG/) WANG H.

Wang H.; PI

WPI: 2004-467656/44.

PT Nanopore analysis system for analyzing polynucleotides includes nanopore
PT device having structure with aperture, and nanopore data analysis system
PT that can generate nanopore data points corresponding to each target and
PT non-target polymer.

PS Disclosure; SEQ ID NO 1; 17pp; English

The invention relates to a nanopore analysis system that comprises a nanopore device and a nanopore data analysis system. The nanopore device has a structure having an aperture. The nanopore data analysis system can generate nanopore data points corresponding to each target polymer and each non-target polymer traversing the aperture; form a distribution pattern of the data points; and analyse a distribution of target polymer data points in the distribution pattern. The distribution pattern includes data cluster(s). The nanopore data analysis system is scored on a computer-readable medium. A mechanism is provided for analysing the distribution of target polynucleotide data points in the distribution pattern. The nanopore data analysis system can also analyse the distribution of the non-target polynucleotide data points. It can also determine a ratio between the target polynucleotide data points and the non-target polynucleotide data points. It can further analyse the distribution of target polynucleotide data points between the two data clusters; compare the distribution of the target polynucleotide data points between the two data clusters to a phosphorylation state standard distribution; and determine a ratio of phosphorylated target

CC polynucleotide to non-phosphorylated target polynucleotides. It can also
 CC determine a cluster score for the target polynucleotide data points in a
 CC defined area; and compare the cluster score for the target polynucleotide
 CC data points to a cluster score for a chemical integrity standard density
 CC distribution for the defined area in a distribution of a target
 CC polynucleotide standard. The invention is useful for analysing polymers,
 CC such as polynucleotides. The inventive nanopore analysis system
 CC potentially provides high speed sampling with single-molecule resolution,
 CC which may enable unprecedented dynamic range and sensitivity in analysis
 CC of samples containing charged polymers such as polynucleotides or
 CC polypeptides. The present sequence is a synthetic construct used in the
 CC invention.

CC Sequence 1300 BP; 1300 A; 0 C; 0 G; 0 T; 0 U; 0 Other;

Query Match 0.8%; Score 80.4; DB 12; Length 1300;
 Beest Local Similarity 41.4%; Pred. No. 1.5e-05;
 Matches 537; Conservative 0; Mismatches 761; Indels 0; Gaps 0;

QY 8426 AAGATTATGAAATATGATTTACACAAAAAATATATTTTCAGAAAGTCAATGCTTGTGCTT 8485
 DB 1 AA 60
 QY 8486 TAAATGATACAAAGATGAAGCAATATGAAATCTTTAGCGGTAGCCGGTGCATGCAC 8545
 DB 61 AA 120
 QY 8546 AAGAAACAACAAACATTTACAGATCAACAAGTTAACTTCAACTGTAATGAGAG 8605
 DB 121 AA 180
 QY 8606 GAAACGTATCTCACTTCGTGCAGAAAGCTTTGGCTAAATAATGAAATATGAAATGTAA 8665
 DB 181 AA 240
 QY 8666 AAGGAATCTGAGAGAGCTTTAGTCGAGCGGAAACGCGCCGTTGAAAAATTAACAAGA 8725
 DB 241 AA 300
 QY 8726 GTACTACAGAGCATTTGTTGTCAGAGAAATTGGGAAATTGGAGATTAATTAGAAAGATTG 8785
 DB 301 AA 360
 QY 8786 CAAGAGATATATCGAATTGTATAGATCAACGAGACGAAACCAAGAGAGCTTTGTCGAA 8845
 DB 361 AA 420
 QY 8846 AGAATGTAATTTCTGTGAAAAATTAACAATTTCAAGGAGAAACAAATATCATTCATTGAAGTA 8905
 DB 421 AA 480
 QY 8906 AAGCAGAAATTTGTGAACCCGGAAGTGTAAATGTATGATGCTTGAATAACTGTATGTAG 8965
 DB 481 AA 540
 QY 8966 ATCTCAAGAAAAAAGTGTGCTATGTGTAATTTGTATTTGGAATGTGATGTAATA 9025
 DB 541 AA 600
 QY 9026 ATGTGATTTAGAAAAATTTAGAACCCAAATTCGAAAGACATGTAATTGTAGAACTACTG 9085
 DB 601 AA 660
 QY 9086 GAAAAACAAGAAATATCAAGATTTTACAGAGCAAAAGTAAATTTCTTGAAAAAGAGACG 9145
 DB 661 AA 720
 QY 9146 CTGCACTGCTGACGCTGCAATATGCAATGTATTCATTTCAATGAGATGATATTAATAATT 9205
 DB 721 AA 780
 QY 9206 TGGCAAGCAGTATGATCTTCTCAATTAATATCAAAAAATTCAAAAATAATATTAATCTT 9265

DB 781 AA 840
 QY 9266 TAGCATCAAGTAGTGAATCGAATGTGATGTTTCATGGGCTGCTGAAGCAAGAGTGCAG 9325
 DB 841 AA 900
 QY 9326 GAGCCAAAGCAGACTAGTGTAAAGAAATCAATTAATAGAACTAATTAATGTTGATTTAG 9385
 DB 901 AA 960
 QY 9386 CAGGAAAAATTTAAACAGAGGAAACATCAATGTATATGCCGATATGATTAATAATA 9445
 DB 961 AA 1020
 QY 9446 ATATAGTATGACAAATTTCTAAGGCTATTGCGGATGCCAAAGTCATGCTGCTGCTT 9505
 DB 1021 AA 1080
 QY 9506 CGGCAACTGCCCATTTGAAAAAAATGAAAGTAAATTTAATATGCGATCCGAGAAATTTA 9565
 DB 1081 AA 1140
 QY 9566 AAAATTAATCTGCAAGATTGGAAGGAAAGCTAATTAATAAAACGTCGTAGATCTAATC 9625
 DB 1141 AA 1200
 QY 9626 AGTATGACTGGTATACGGATTAATATATATATGATGCACTTCTTGAAAAAGCATACAAAAT 9685
 DB 1201 AA 1260
 QY 9686 TGACATATCAATCAAGAGAGAGAAAAAGGAAAAA 9723
 DB 1261 AA 1298

RESULT 16

ADA89806/C
 ID ADA89806 standard; DNA; 4965 BP.

ADA89806;

20-NOV-2003 (first entry)

DE Staphylococcus aureus antigenic partial gene sequence locus 35B.

XX antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
 XX antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
 XX antitumor; immunostimulant; ophthalmological; pathogenic microbe;
 XX bacteremia; septic shock; organ infection; skin infection;
 XX bacterial basal colonisation; bacterial eye infection; septicemia;
 XX tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
 XX sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
 XX necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
 XX gastro-enteritis; dysentery; shigellosis; skin disorder; gene; ds.

OS Staphylococcus aureus.

PN WO2003011899-A2.

PD 13-FEB-2003.

PF 02-AUG-2002; 2002WO-GB003606.

PR 02-AUG-2001; 2001GB-00018825.

PR 09-JAN-2002; 2002GB-00000349.

PA (UTSH-) UNIV SHEFFIELD.

PA (BIOS-) BIOSYNEXUS INC.

PI Foster S, Mond J, Clarke S, McDowell P, Brummel K,

DR WPI; 2003-256434/25.

PT New antigenic polypeptides from *Staphylococcus aureus* or *S. epidermidis*, useful as a vaccine for immunizing humans against e.g. bacteremia, septic shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhea or impetigo.

Claim 1; Page 68-69; 189pp; English..

The present invention describes an antigenic protein or its part, which is for use as a vaccine. The antigenic protein is encoded by an isolated DNA molecule of *Staphylococcus aureus* or *S. epidermidis* partial gene sequences (designated dnaSA and dna SE, respectively; and which encodes a protein expressed by a pathogenic organism. Also described: (1) a vaccine composition comprising at least one antigenic protein; (2) a method of immunising an animal against a disease or condition caused by a pathogenic microbe by administering the antigenic protein or the vaccine; (3) an antibody or its binding part obtainable by the method above; (4) preparing a hybridoma cell line producing monoclonal antibodies; (5) a hybridoma cell line produced by the method of (4); and (6) identifying opsonic antigens expressed by a pathogenic microbe. The antigenic proteins have antibacterial, neuroprotective, immunosuppressive, antiinflammatory, antituber, immunostimulant and ophthalmological activities, and can be used in vaccines. The antigenic proteins or vaccines can be used for immunising an animal (specifically a human) against a disease or condition caused by a pathogenic microbe, e.g. bacteraemia, septic shock, organ infection, skin infection, bacterial baasal colonisation, bacterial eye infections, septicæmia, tuberculosis, bacteria-associated food poisoning, blood infections, peritonitis, endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea, strep throat, streptococcal-associated toxic shock, necrotising fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis, dysentery, shigellosis, *S. aureus*-associated septicæmia, food-poisoning, skin disorders, *S. epidermidis*-associated septicæmia, peritonitis or endocarditis. The present sequence represents a *S. aureus* partial gene sequence from the present invention.

Sequence 4965 BP; 1119 A; 870 C; 861 G; 2115 T; 0 U; 0 Other;

Query Match 0.8%; Score 79.2; DB 8; Length 4965;

QY	3782	TTCAAAGTAAACAAGTCCTTGGTTAAAAATAGTATTTCCGAATGCCATTAATTTA	3641
Db	4395	TTAATATAAAAAGCAACTGCACGTATATGAATACAGCAATTTTAAATACAAATTGCAAG	4336
QY	3842	AAGTAATAGCTTTAGTGGAGAACTCAAGTAGCAGCAGACAGTCTTGGAAACAGTTA	3901
Db	4335	AGATTCAAGCTACGCCAGATGCACAGATGAAGAAAAACAAGCATGTATGCTAACCA	4276
QY	3902	AAGAAAGTGAAGACAAAGAAAAATTATCTATTGGAACTTCTGCTTATCAACTTAG	3961
Db	4275	ATACTGAATAATGTGTAAAGCAAAATCAAGCCATTTGACGACAACTATACGCACAAAGTTG	4216
QY	3962	TGAACAAATGAATTTCTGCATAATTCAGAAAAATTAATACGTAGCAGAGAAATCTGAAAGCC	4021
Db	4215	ATGAAGCTTAAAGCAAAATTCAGAAAGAGGATTAATAGCCGTAAACCAAAATTTGTAAAG	4156
QY	4022	AAAAAATGAGTATTGATGTCATCTGCTTATCAAGCGCACCCCAAGTGCAGAGACTTTAA	4081
Db	4155	AACAAGCGCTTAANATGATGAATTAATGATCAATTAACAAGCAACCAAAATGTTATCATATA	4096
QY	4082	ATTTAACAAGCTGAGAAAGTCAATGGAATCTTAGGGGCTACTGTGACTGTTGCCAAATTA	4141
Db	4095	ATGATTCAGAACCCCTTCAACAGAAAGAAAAAGAACGCTATTTCACAAATTAACCAAGCAG	4036
QY	4142	ACAAACAAGTAATGCTCTTATTAGTGGGAGATATACTAACGTTATGAGCGACG	4201
Db	4035	TTACAGACCGCAAAAATATATATTAACAGCTGCAACTGATGATATATGTTAGTCAAGCGGA	3976
QY	4202	CAAAAGCTCTTTTACCAACCACTCAAGTGAAGTCTGCAAGTACGACGGAGGAGCAATTA	4261
Db	3975	AAGACGCTGGAAGATTCATTTCAAAGCACCACCAACGCAACAGCGTTAAATCAATG	3916

QY	4262	GTCTGAGCGGGATTAGAAATTATCAGGGCGCTTCTCTGCAATAAGTTGACAATG	4321
Db	3915	CTAAAAATGATGTTGATCAACGCTGTGACAATCMAATAGCAATGTATATACAAGT	3856
QY	4322	ACGTGAAGCTAGCGTTGATTAATCTTCCATCGAGAGCTAATGAATCAATGTGATG	4381
Db	3855	GTGCTACACTGGAAGAAAAATGCAAAAAGATTATGTTTAAAGCTAAAGAAAAAG	3796
QY	4382	CCAAAGATGTCAAGAAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGAA	4441
Db	3795	CGTATCAAGATATCTTAAATGCAACAACATATATGTTAGCAAAATTAAAGATCAAG	3736
QY	4442	AAGATAAAAAATTTTGAAGATCGTGTATTAATACAGCTGAAATGCTTATATACGA	4501
Db	3735	CAGTTGCTGATATTCAGAGTATTACTGCAGATCAACAATTAAGATGTTG-----CGA	3682
QY	4502	AGGAACAATCTGAAAAAGCAAAAGAAAAAAGAGAGCGCATGTGTAATGCTGCTTAT	4561
Db	3681	AAGATGAATTAAGCAACAAAGCAACGAAACAAAAGGCGTTATTATGCAAACTGCAGATG	3622
QY	4562	CGGTTGCTGGAACGGATTAATCCGCTGGAGAGTAGCTATTATAGTCAATACTGTTAAAA	4621
Db	3621	CGACTACTGAAGAAAAAGAACACAGCAATCAACAGTAGACGCAATTAACACAAGSTA	3562
QY	4622	ATTAATTTAAAGCAAAATTAGTGAAGCAATAGCAAGACCGAGAGAGATAAATTCATG	4681
Db	3561	ATCAAAATATTGAATAATGCAAGTCAATCATATGATATTAACATGCAAAAAGATTAAGCA	3502
QY	4682	CGAAACATGTAAATGTGAGGCAAAATCATCTACTGTGTGTGAATGGCGCTTCTGAC	4741
Db	3501	TTCAAGCAATTGACCAATTCACATCAACAGATGTTAAACGAATGCAAGAC--GGAA	3443
QY	4742	TTGCTATCAGAAAGATGCTTTTCAGAAATGGGATGTGAGCATGCAAGACTTATCA	4801
Db	3442	TTGCTACTGAATGCAAAATTAATATATCTGAAATACTTATATATATAGAGACT-----	3389
QY	4802	ATGACAGCATTGCAAGGTGGATTAAGAAAGAAATTTCTGCTGATTTCTTAATGTGACG	4861
Db	3388	-----ACTTAATGAAGAAAAAGGTAAACGATATGTGACAGTTAGAGCAATATGAAGAG	3334
QY	4862	CAATTAATTCATTCTTGGGCTGAATGTGGCGGAACCATTCGCGGTTCTTTCTTAAGG	4921
Db	3333	GTTTAAATTAATTAATTAATGACCAACTCTACAGGTGATGTACTACGCTTAAGATACAG	3274
QY	4922	CGGTGAGAGCTGCTTTGCGAATATATCTCTCAATTAATAAACCCTGCTTGTGATTAAG	4981
Db	3273	CAGTACAAAAGTTCAACACTCTATGCAATCTCTGTTAAGAAACAGCAGGTAAAAAG	3214
QY	4982	GAAACAAAGTAAATCTTTTATGAGAAAGATACMAAAGTCAATGTACAAGCTTTGATG	5041
Db	3213	AATTGATCAAGCTGCAGCTATTAAGAAAAACAAATATGAACAAACCAATGCATCAC	3154
QY	5042	ATTCTCATTTTCAAAAGCTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGAATCG	5101
Db	3153	AACAAGAAATTAATGATCAAAACAAAGAAAGTTGTACTGAAATTAATCAAGCGAAAACAA	3094
QY	5102	GAGGAATGTAATCTGTCAATGCTGTTCTGATGAAACGGAAGCTTTGTGATGATCTTG	5161
Db	3093	ATGTGATCAATCATCAACAAATGAATATGTTGATATATGACGTTTAAAGAGAAACCTA	3034
QY	5162	AGTTTGAAGAGTAAAGTTCTTTCATATGTAGTCAAAAGATCAAAAAACATAAATACAA	5221
Db	3033	AAATTATATGCAGTTAAACATTTAGTGATCAAAAAAGATGCTTTAGCTAAATTTGAAG	2974
QY	5222	TTGCGGAAATGCAAAATGAGAAAAAGCGCTGAGTGTGAGCAACAGTCTCATACAA	5281
Db	2973	ATGCAATATATGCTAAAGTAAACGAAGGGAATTACTTAAGCATGCACTTCAAGTAA	2914
QY	5282	ATATTGGAACAAATCAGTTATAGTATGTATGTAATAAACAATTAATTAACAACGCGAATG	5341
Db	2913	TTGCTGAAGCGAAACAAAACCTTGCTGAATTTAAACAAACCTGCGATCAAAATGTTAATC	2854
QY	5342	ATCAAGATGAATAAATATCAATGTGATGCAAAAGATTATCTATGCAATATCTATAG	5401

```

Db      2853  AAGTACTCTTAAAGATGACATTGAAGTTCAAAATTCATTAATGACTTAGATAATATTAAACG 2794
Oy      5402  CAGTCGAGATTGGAGAGCAAAAAGAGCCTGTGTGCAAGAGCTTCGCAAGTACTACT 5461
Db      2793  ATTACAAATTCACACAGGTAAAGAAAGATCAGTACACACATTTATATGTGTTAGAG 2734
Oy      5462  TGAATTAAGCAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAAGATTAGAGAG 5521
Db      2723  ATCAGAAAGAAAAATATATTTCAGCTGACACTTAATGCAACAGAAATGAAAAAGCAAG 2674
Oy      5522  AAAATTAATGGAATTAAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATACGAGTCAAG 5581
Db      2673  CAATTAGCAGATTACCAAAATG---TTCAAACCTGCATTAGAAAAGCATTTAATTAATGTG 2617
Oy      5582  TGGTCACAAATGCGACAGTGTCTTCGCGAGCAAGTGAACAAGCTCAGTAGAGAGTGGAG 5641
Db      2616  TGGATTAATGTTAGAGCTGTGATGATGATTAAACAAAGTAAGCAGCAATGTATGCTATTC 2557
Oy      5642  TAGCAGTTAATTAATTAACACAAATTAATCTTGCACATTAATAAATAGTACTCAAAATG 5701
Db      2556  AAGTAGATGCTACTGTAACTTAAGCAAGCAACCAAGCTATTGAAGTTAAAGCAGAAAGATA 2497
Oy      5702  TAGCAATGCTTTGTTAAAAAG-----CAAATCTCATTCATCTTA 5740
Db      2496  CGAAAGAAATCTATTGATCAAAAGTGAACCAAGTTAATGCTGAAGAAAAAAGTGAAGCATTTAG 2437
Oy      5741  TTAATAACAATTGGAGCTGGAGAGTTGAGCTGGAGAGAGCTGGAGAGTGAAGTTCG 5800
Db      2436  CAATGATTAATTAACAAATTAACAGATCAAGTAAACAAAGTATTAATCTATGCAACACACTG 2377
Oy      5801  TAGCAGTAATTAAGATTGTAATTAATTAAGATAGCAGAAATTAATGATCAAAATCAGCTG 5860
Db      2376  CTGAAGTTGAAAAGCAAGCAAGCTCAAGGACTTGAAGCAATTTGATTAACATTCAGTCAACT 2317
Oy      5861  CGAAGGAAATGTCGAGATTATTACAGAGTCTGATCGGTAAATTCCTAATTAATGCAAGAA 5920
Db      2316  CAACGAAAAACMAAAAGCTATCGAAGTAATGAAAGTCACTAGACCAAGATTGAAGCAG 2257
Oy      5921  CAGTGTGAGAGTGGCCCGTGCAGCAATAGAGCCTCAACAGCTGATGAATGAATTAAG 5980
Db      2256  GTGTAATATGTCACAGCTGATGTCTACACTGAAGAAAAAGAACGTTTACGAATGCTTTAG 2197
Oy      5981  GATCTACAAAGCATATGTAAAGATTCTACAGTATGCTTAAAGAGAAACAGATGAT 6040
Db      2196  AAGACATTTTATCAAAAGCACTGAAGATTTTCTGATGAACCTAACAATGAGAAATCG 2137
Oy      6041  ATATTACTACTCAAGGCAAGTAGATTAAGTGTGATTAAGTATTCAAAAATCTTAATA 6100
Db      2136  CTACTGTCAAAAATAGTGGCTTGAACAACCTTAAGCAACAGTATTAATCCTGAAGTTA 2077
Oy      6101  TTAAGAGAGACTTATCACAAAAAAGAAAAAATAAGTATTAATAAAGATTTGTTACCAATA 6160
Db      2076  AGAAAAATGCTTTGAGAGCAATCAGAGAAAGTGTTTAACAGCAAAATAGAAAAATTTAAAA 2017
Oy      6161  GTTCAGTACTACTACTT 6178
Db      2016  ATGCAGATGCAGATGCAT 1999

RESULT 17
AAS52179
ID      AAS52179 standard; DNA; 7434 BP.
XX
AC      AAS52179;
XX
DT      13-FEB-2002 (first entry)
XX
DE      Staphylococcus aureus DNA for cellular proliferation protein #596.
XX
KM      Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX      antibacterial; drug design.
XX

```

```

OS      Staphylococcus aureus.
XX
PN      MO200170955-A2.
XX
PD      27-SEP-2001.
XX
PF      21-MAR-2001; 2001WO-US009180.
XX
PR      21-MAR-2000; 2000US-0191078P.
PR      23-MAY-2000; 2000US-0206848P.
PR      26-MAY-2000; 2000US-0207727P.
PR      23-OCT-2000; 2000US-0242578P.
PR      27-NOV-2000; 2000US-0253625P.
PR      22-DEC-2000; 2000US-0257931P.
PR      16-FEB-2001; 2001US-0269308P.
XX
PA      (ELIT-) ELITRA PHARM INC.
XX
PI      Haeelbeck R, Ohlsen KU, Zyckind JW, Wall D, Trawick JD, Carr GJ,
PI      Yamamoto RT, Xu HH;
XX
DR      WPI; 2001-611495/70.
DR      P-P9DB; AAU34320.
XX
PT      New polynucleotides for the identification and development of
XX      antibiotics, comprise sequences of antisense nucleic acids.
XX
PS      Claim 27, SEQ ID NO 4761; 511pp; English.
XX
CC      The invention relates to antisense inhibitors of gene essential to
CC      prokaryotic cellular proliferation, their use in identifying the genes,
CC      their use in the discovery of novel antibiotics, the essential genes
CC      themselves and the encoded proteins. The prokaryotes used are Bacillus
CC      coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC      Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC      useful for the identification of potential new targets for antibiotic
CC      development. The antisense nucleic acids can also be used to identify
CC      proteins used in proliferation, to express these proteins, and to obtain
CC      antibodies capable of binding to the expressed proteins. The proteins can
CC      be used to screen compounds in rational drug discovery programmes. The
CC      antisense nucleic acid sequence is also useful to screen for homologous
CC      nucleic acids which are required for cell proliferation in a wide variety
CC      of organisms. The present sequence encodes an essential prokaryotic
CC      cellular proliferation protein. Note: The sequence data for this patent
CC      did not form part of the printed specification, but was obtained in
CC      electronic format directly from WIPO at
XX      ftp.wipo.int/pub/published_pcc_sequences
XX
SQ      Sequence 7434 BP; 3197 A; 1326 C; 1298 G; 1613 T; 0 U; 0 Other;
XX
Query Match      0.8%; Score 79.2; DB 4; Length 7434;
Beet Local Similarity 41.3%; Pred. No. 4.4e-05;
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;
Oy      3782  TTCAAAGTAAACAAGTCTTGTGTTAAAAATAGATATTGAAATGCCAATAATTTA 3841
Db      3659  TTAATTAATAAAGCAACTGACGTAATGAAATCAACGCAATTTTAATAACAATGCAAG 3718
Oy      3842  AAGTAATGCTTTGAGTGAAGAACTCAAGTACGACGAGCAGAGTTTGAAGCAGTTA 3901
Db      3719  AGATTCAAGCTACCCAGATGCAACAGATGAAGAAAAACAAGCAGCTGATCTGAAGCA 3778
Oy      3902  AAGAAAGTGAAGGACAAGGAAAAAGTTATCTAATGGGAAGCTTCGCTCTATCAACTAG 3961
Db      3779  ATACTGAAAATGTTAAGCAATTCAGGCAATTTAGCGCAACTAGTAAAGCAAGTTG 3838
Oy      3962  TGAACAATGAAGTTTTCGCAAAATCAGAAATTAATATACAGTACGAGAGAACTTGAAGCC 4021
Db      3839  ATGAAGCTTAAGCAAAATGCAAGCAGCAAGTAAATGCGGTATACCAACAAAGTGTGAAGA 3898
Oy      4022  AAAAATGATGTGATGTCCTGCTTATCAAGCGGACCAACCAAGTGCAGAGCTTTAA 4081
Db      3899  AACCAAGCGCTTAAGATGAATGATCAATTATCAAGCAACCAACCAATGTTATCAATA 3958

```

4082 ATTTACAAAGCTGAAAGTGAATGCACTGTAGGGGCTACTGTGACTGTGCGCAATTAA 4141
4112 ACAACAAAGTAATGCTTATTAGTGGGAGATATCTAAAGTTATCGAGCGGAGC 4201
4019 TTACAGACGCGAAAATTAATTTACAGCTGCACATGATATATGTGTAGATCAGGCGA 4078
4202 CAAAAGCTCTTTAGCAACCACTCAAGTGTGCTGCACTGACGAGCGGAGCAATTA 4261
4079 AAGAGCTGGAAGAAATTCATTCAAGAGCAGCAACAGCAGCGGTTAATCAAAAG 4138
4262 GTTCTGAGACGGGATTAGAAATTTATCAAGGGGCTGTTCTGTCAATTAAGTTGACATG 4321
4139 CTAAAAATGATGTGTGATCAAGCTGTGACAACTCAAAATCAAGCAATGTATTAACAATG 4198
4322 AGGTGAGAGCTAGCGTTGATTAATCTTCATGGAAGAGCTAATGAATCAATGTGCTAG 4381
4199 GTGCTACAACTGAAGAAAAATGACGCAAAAGATTTAGTTTAAAGCTTAAGAAAAAG 4258
4382 CCAAAAGTGTCAAGGAGAGTTCGATCTAGCAAAAGAAATATCAGGCTTTACTAATGAA 4441
4259 CGTATCAAGATATCTTAATGCAACAACATTAATGATTAAGCAAAATTAAGATCAAG 4318
4442 AAGATAAAAATATTTAGAAAGATCGTGTATTAATACGACTGGAATGTTATTAATGCA 4501
4319 CAGTGTCTGATATTCAGGGTATTAATCTGCAAGATACAACTAATAAGATGTTG-----CGA 4372
4502 AAGAACACTAGAAAAAGCAAAAGAAAAAGAGACGGTCAATTTGTAATGCTGCTTAT 4561
4373 AAGATGAATTAAGCAAAAAGCAAAAGCAAAAGCGCTATTTGCAAACTGCGAGATG 4432
4562 CGGTTCTGGAACGAGTAAATCCGCTGAGAGTAGCTATTTGCAATTAATCTGTTAAA 4621
4433 CGACTACTGAAGAAAAAGAACGAAATCAACATTAAGACGCAATTAACGAAAGTA 4492
4622 ATTAATTTAAAGCAATGAGTGAAGCAATTAAGAAAGCCGAGAGAGATTAATTCATG 4681
4493 ATCAAAATATTTGAATAATGCAAGTCAATCGATGATGTAACACTGCAAAAGATTAAGCA 4552
4682 CGAAATGTAAATGTGAGGCGAAATCATTCTGTTGTTGTAATGCGGCTTCGAGC 4741
4553 TTCACCAATGACCCCAATTCAGATCAACAGATGTTAAACGAATGCAAGAC-----GGA 4611
4742 TTGCTATCAGCAAGATGCTTTTTCAGAAATGGGATCTGGAGCATGGCACTTATCA 4801
4612 TTGCTAATGAAATGCAAAATTAATAATTAATGAAATCTTAATTAATGAGACT----- 4665
4802 ATGACACGATTTGCAAGGTGATTAAGAAAGAAATTTCTGCTGATTCCTTAATGTGAACG 4861
4666 -----ACTAATGAAGAAAAAGTACGATATGTGACACAGTTAGAGCAGCATATGAAGAG 4720
4862 CAAATTAATTCATTTCTGGGGTGAATGTTGCGGGAACATTTGCCGGTCTCTTTCTAAGC 4921
4721 GTTTAATATATTAATTAATGACAACTAATACAGGTGATTAATTAATGATTAATGATTAAT 4780
4922 CGGTGAGAGCTGCTTTTTCGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4981
4781 CAGTACAAAAGTTCAACAATCTCATGCAATCTGTTAAGAAACGACGAGTAAAAAG 4840
4982 GAAAGAGGTAAATCTTTTATGAGAAAGATACAAAGTCAATGATCAAGGCTTTGATG 5041
4841 AATTAAGTCAAGCTGAGTGAATGAAGAAACAAATTAAGAAACCAAAATGATCATC 4900
5042 ATTCTCATTTACAAAGTTTCTGCTGAGCGCTGCAAGTATTAAGCAGCTGGAATCG 5101
4901 AACAGAAATTAATGATGCAAAACAAAGAGTGTATGATTAATTAATTAATTAATTAATTAAT 4960
5102 GAGGAATGATCTGATCAATCGTGTCTGATGAAGCAAGGCTTTAGTATGATCTG 5161
4961 ATGTGATCAATCATCAACAAATGAATATGTTGATTAATGCAATTAAGAAAGAAAGCTA 5020

5162 AGTTGAAGAGTAGTACTTTCAATGTAGATGCAAAAGATCAAAAAACAATTAATCA 5221
5021 AATTAATGACATTTAAACATTTAGTGTACAAAAAGATCTTACTTAATTAAG 5080
5222 TTGCGGAATGCAAAATGAGAAAAAGCGCTGAGGTTGAGCAACAGTTGCTATCA 5281
5081 ATGCAATTAATGCTAAAGTAAAGAAAGGAGATTAATTAAGCATTCAGCTCAAGTGA 5140
5282 ATATTGAAAACATCACTGATTAAGTATTTGTAATAAAACGTAATAATTAACGGCGAATG 5341
5141 TTGCTGAAGCGAAACAAAACCTGCTGAATTAATAAAACAACTGCGAATCAAAATGTTAAATC 5200
5342 ATCAAGATAGAAAAATATCAATGTGACTGCAAAAGATTAATTAATTAATTAATTAATTAAT 5401
5201 AAGCTACTTCTAAGATGACATTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5260
5402 CAGTGGAGTTGAGAGAGCAAAAGAGCGCTGTGCAAGAGCTTCTGCAATTAATTAATTAAT 5461
5261 ATTACCAATTCACACAGGTAAAGAAATGATCAACAGATTAATTAATTAATTAATTAATTAAT 5320
5462 TGAATTAAGACATTTCTTCTCATGTTGATCAAACTGATTAATTAACAAAGATTAAGAGAA 5521
5321 ATCAGAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5380
5522 AAAAAATGGAATAAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATTAAGATCAAG 5581
5381 CAATTAAGCAATTTGACCAAAATG-----TTCAAACTGCTATTAAGAAATTAATTAATTAAT 5437
5582 TGATCAAAATGCAAGTGTCTTCCGAGCAAGTGAAGCAAGCTGCAAGTGAAGCTGAG 5641
5438 TGATTAATGATGATGATGATGATGATGATTAATTAACAAAGTAAAGAGCAATTAATTAATTAAT 5497
5642 TAGCAGTTAATAAATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5701
5498 AAGTATGATCTACTGTTAAACCTTAAGCAAGCAAGCTATTAAGTAAAGCAAGAAATA 5557
5702 TAGAAATGCTTTGTAAGAAAG-----CAAAATGCTCAATCTTA 5740
5558 CAAAGAAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5617
5741 TTAAAACAATTTGAAATTTGAGCTGAGTGTGAGCTGAGAGCTGAGAGTGAAGCTGTTCTG 5800
5618 CAATGATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5677
5801 TAGCAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5860
5678 CTGAAGTTGAAAAAGCAAGCTCAAGAGCTTGAAGCATTTGATTAATTAATTAATTAATTAAT 5737
5861 CGAAGGAATGTCGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5920
5738 CAAAGAAAAACAAAAGCTATGAGAAATTAAGAACTGCACTAGACCAAGTTGAAGCAG 5797
5921 CAGTGTGAGAGTGGCCGCTGAGCAATAGAGCTCAACAGTGTGAATGAATTAATTAATTAATTAAT 5980
5798 GTGTAATGTCACAGCTGATGTCACAACTGAGAAAAAGAGGTTTAAGAAATGCTTTAG 5857
5981 GATCTCAAAAGCATATGTAAGATTTCTACAGTGTGCTTAAGAAAGAAAGCAAGATTAAT 6040
5858 AAGACATTTTATCAAAAGCAATGAGATATTTCTGATCAAACTGCAAAATGCGAAATCG 5917
6041 ATTTACTACTCAAGGCAAGTGAATTAAGTGTATTAATTAATTAATTAATTAATTAATTAATTAAT 6100
5918 CTACTGTCAAAATTAATGAGCTTGAACCACTTAAGCAACAGCTATTAATTAATTAATTAATTAAT 5977
6101 TTAAAGAACTATCACAAAAAAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6160
5978 AGAAAAATGCTTTGGAAGCAATTAAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 6037
6161 GTTCACTACTCACTACT 6178
6038 ATGCAGATGCAAGATGAT 6055

RESULT 18
AB114965
ID AB114965 standard; DNA; 7434 BP.
XX
AC AB114965;
XX
DT 06-MAR-2003 (first entry)
XX
DE Pathogen specific antigen related staphylococcal DNA SEQ ID No 201.
XX
KM Antibacterial; virulence; fungicide; protozoasicide; cytostatic; anti-HIV;
KM hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KM auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KM autoimmune disease; HIV; hepatitis; gene; ds.
XX
OS Staphylococcus sp.
XX
PN MO200259148-A2.
XX
PD 01-AUG-2002.
XX
XX 21-JAN-2002; 2002WO-EP000546.
XX
PR 26-JAN-2001; 2001AT-00000130.
XX
PA (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX
PI Meinke A, Nagy E, Von Ahnen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vytvytska O, Elz H, Dryla A, Weichardt T, Hafner M;
PI Tempelmeier B;
XX
DR WPI; 2003-075410/07.
XX
PT Identifying, isolating and producing hyperimmune serum-reactive antigens
PT from a pathogen, for preparing vaccine or medicament for treating or
PT preventing e.g. staphylococcal infections, comprises providing antibody
PT preparation.
XX
XX Example 7; Page 179-180; 252pp; English.
XX
PS The invention relates to a novel method for identifying, isolating and
XX producing hyperimmune serum-reactive antigens from a pathogen, tumour,
CC allergen, a tissue or host prone to auto-immunity, where the antigens are
CC used in a vaccine, comprises providing antibody preparation from a plasma
CC pool of a type of animal, or individual sera with antibodies against the
CC specific pathogen, tumour, allergen, tissue or host prone to auto-
CC immunity. The hyperimmune serum-reactive antigens comprising any of the
CC 62 sequences of 53-2261 amino acids fully defined in the specification,
CC or their hyperimmune fragments are useful for the manufacture of a
CC pharmaceutical preparation, particularly a vaccine against staphylococcal
CC infections or colonisation against *S. aureus* or *S. epidermidis*. The
CC preparation of antibodies is useful for the manufacture of a medicament
CC for treating or preventing staphylococcal infections or colonisation
CC against *S. aureus* or *S. epidermidis*. The antibody preparations may also
CC be used for diagnostic and imaging purposes. Other conditions that can be
CC treated include cancer, autoimmune diseases or infections caused by viral
CC (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
CC polynucleotide sequence represents staphylococcal DNA relating to the
CC method for identifying and producing pathogen specific antigens of the
CC invention
XX
SQ Sequence 7434 BP; 3196 A; 1327 C; 1298 G; 1613 T; 0 U; 0 Other;
XX
Query Match 0.8%; Score 79.2; DB 8; Length 7434;
Best Local Similarity 41.3%; Pred. No. 4; e-05;
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

Db 3719 AGATTCAAGCTACGCCAGATGCCAACAGTGAAGAAAAACAAGCTGATCTGAAGCA 3778
Qy 3902 AAGAAAGTGAGAGCAAGAGAAAAAGTTATCTATTGGGAACTTGCTTCTATCAACTTAG 3961
Db 3779 ATACTGAAAAATGTTAAAGCAAAATCAAGCATTTTCAGACGCAACTACTAACCCACAAGTTG 3838
Qy 3962 TGAACAAATGAAGTTCTGCAAAATCAGAAAATTAATACAGTAGCAGAGAACTTGAAAGCC 4021
Db 3839 ATGAAGCTAAAGCAAAATGCAGACAGCATTAATGCGGTAAACCAAAAGTTGGAAAGA 3898
Qy 4022 AAAAAATGAGATGTGATCTACTGCTTATCAAGCGGACACCCAGAGTGACAGAGCTTTAA 4081
Db 3899 AACCAAGCGGCTPAAAGATTAATTAATCAATTACAAAGCAAGCCAAACAAATGTATCAATA 3958
Qy 4082 ATTACCAAGCTGAAAGTCAATGAACTGTAGGGGCTACTGTGACTGTGCCAATTTAA 4141
Db 3959 ATGATCAGAACGCTACCAACACAGAAAAAGAAAGCAGCTATTCCAAATTACACACAGAG 4018
Qy 4142 ACAACAAAGTAATGCTTCTATTAGTGTGGAGATTAATCTAACGTTAATCGAGCGAGCG 4201
Db 4019 TTACAGACGCGAAAAATTAATTAATTAACAGCTGCAACTGATGATTAATGTTAGATCAGCGCA 4078
Qy 4202 CAAAGCTCTTTTAGCAACCACTCAAGTACTGCTGCAAGTACGACCGGAGAGGACAAATTA 4261
Db 4079 AAGAGCTGGAAGAAATTCATTAACAGACGCAACGCAACAGCGGTTAATCAATTAATG 4138
Qy 4262 GTTCGAGCGGAGATTAGAAATTTATCAAGGGGCTGTTCTGCAATPAGATTGACAAATG 4321
Db 4139 CTAAAAATGATGTTGATCACTACCTGTGACACTCAAAATCAAGCAATTGATATCAACTG 4198
Qy 4322 ACGTGAAGCTAGCGTTGATTAATCTTCATCGAAAGAGCTTAATGAATCAATGTCATTG 4381
Db 4199 GTGCTACCAACTGAAGAGAAAAATGACGCAAAAGATTTGTTTAAAGCTTAAAGAAAAAG 4258
Qy 4382 CCAAAAGTGTAAAGAAAGTCTGATCTAGCAAAAGAAATTCAGGCTTTACTAATATGAA 4441
Db 4259 CGTATCAAGATATCTTAAATGCACAAACACTTAATGATGATCGAAATTAAGATCAAG 4318
Qy 4442 AAGATTAATAATTTTGAAGATCGTGATTTAATATACACTGAAATGGTTATTATACGA 4501
Db 4319 CAGTTGCTGATATTCAGAGTATTACTGCAGATACAACTTAATTAAGATTTG-----CGA 4372
Qy 4502 AGGAACCACTGAAAAAGCAAAAGAAAAAGAGAGCGGCTATTGTAATGCTGCTTAT 4561
Db 4373 AAGATGAATTTAGCAACAAACAAACGAAACGAAACGCTATTATGCAAACTGACATG 4432
Qy 4562 CGGTTCGTGAGACGATTAATCCCTGAGAGATGATGATTTGCACTCAATCTGTTAAA 4621
Db 4433 CGACTACTGAAGAAAAAGAACAGCAAGCAATCAAGTAGACGCAATTTAACACAGGTA 4492
Qy 4622 ATTAATTTAAGCAAAATTTAGTGAAGCAATTAAGAAAGCCGAGAGGATTAATTTCAATG 4681
Db 4433 ATCAAAATTTGAAAAATGCAAGATCAATCGATGATTAACACTGCCAABAATATGCA 4552
Qy 4682 GAAAACATGAATGTGAGGCAAAATCATCTACTGTTGTGTGAATGCGGCTTGTGAC 4741
Db 4553 TTCAAGCAATTTAGCCCAATTCAGACATCAAGATGTTAAACGATCAAGAGC-GGAA 4611
Qy 4742 TTGCTATGCAAGAAAGTCTTTTCAGAAATGGGATCTGAGCATGCGCAACCTTATCA 4801
Db 4612 TTGCTTAATGAAATGCAAAATTAATACTGAAATTAATTAATGAAGCT----- 4665
Qy 4802 ATGACACGATTTGCAAGGTGATTAAGAAAGAAATTTCTGATGATCTTAAATGTGAAGC 4861
Db 4666 -----ACTAATGAAGAAAAAGATTAACGATTTTGAACGCTTGAACGACATTAAGAG 4720
Qy 4862 CAAATTAATTCATCTTGGGATGAAATGTTGGGGAACCATTTGCGGTTCTCTTTACGG 4921
Db 4721 GTTAAATTAATTAATTAATGACCAACTACTACAGGATGATTAATCTCTAAAGATACAG 4780
Qy 4922 CGGTAGAGCGCTTTTTCGAAATTAATCTTTCATTAATTAACCTCTGCTTTGATTAAG 4981
Db 4781 CAGTACAAAAAGTTCAACACTTATGCAATCTGTTTAAGAAACACAGAGTAAAGAAAG 4840

```

QY 4982 GACGAGATTAATCTTTAGTGGAAAAGATACAAAAGTCAATGTACAGTTTGAATG 5041
DB 4841 AATTAGATCAAGCTGACCTGATTAAGAAAACAAATATGAAACAAACACCAATGCATCAC 4900
QY 5042 ATTCTATATTACAAACGTTTCTGCTGGAGCGCTGCAAGTATTATAGCAGCGTGGATCG 5101
DB 4901 AACAAAGAAATTTATGATGCAAAAACAGAAAGTTGATACGAAATTAATCAGGAAACAA 4960
QY 5102 GAGGAATGGTATCTGTCAATCGTGTCTGTATGAAACGGAAGCTTTAGTTAGTTCTG 5161
DB 4961 ATGTGATCAATCAATCAAAATGAAATGTTGATTAATGAGTTAAAGAGGAAAGCTTA 5020
QY 5162 AGTTGGAAGGAGTAAGTTCTTCAATGTATGATGCAAAAAGATCAAAAACAAATTAATACA 5221
DB 5021 AAATTAATGCGATTAAACATTTAGTGTAGTACAAAAGAAAGATGCTTTAGTAAATTTGAAG 5080
QY 5222 TTGCGGAAATGCAATGAGAGGAAAAGCGCTGAGTGTGAGCAACAGTTGCTCATACAA 5281
DB 5081 ATGCAATTAATGCTTAAAGTAAACGAAAGCGGATTAACCTTAACGATCGACTTAAGTGA 5140
QY 5282 ATATTGAAAAACAAATCAGTTATAGCTATTGTAAAAACAGTAAATTAACACGCGGAATG 5341
DB 5141 TTGCGAAGCGGAAACAAAACCTTGCTGAATTAACAAACCTGCGATCAAAATGTTAAATC 5200
QY 5342 ATCAAGATGAAAAAATATCAATGTGACTGCAAAAAGATTATCTATGACCAATACTATAG 5401
DB 5201 AAGTACTTCTTAAGATGACATTTGAAGTTCAAAATTCATATATCACTTAATTAATTAACG 5260
QY 5402 CAGTGGAGTGTGAGAGGCAAAAAGAGCTCTGTGCAAGAGAGCTTCTGCAATGACTACT 5461
DB 5261 ATTACACAAATTCACAGGTAAGAAAAGAAATCAGCTACACAAATTTATATGCTTAATGCA 5320
QY 5462 TGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTAACAAGATTTAGAGAG 5521
DB 5321 ATCAGAGGAAAATATATATTTCAGCTGACATTAATGCAACACAAATGAAAAACAAACAG 5380
QY 5522 AAAATTAATGAAATTAAGAAAAGCAAAATGTTAATGTTTACTGAAAAATACAGAGTCAAG 5581
DB 5381 CAATTAAAGCAAGTGTGCAAAAATG--TTCAAACTGCAATTAAGAAAGCATTAATTAATG 5437
QY 5582 TGGTCACAAAATGCGACAGTCTTTCGAGCAAGTGGACAACTGAGTAGAGAGCTGAG 5641
DB 5437 TGGATTAATGTAACGTTGATGATGATTAACAAAGTAAAGAGCAATTAATGCTAATTC 5497
QY 5642 TAGCAGTTAATAAATTAACAAAAATTAATCTTGACATTAATAAATACTACTCAAAATG 5701
DB 5497 AAGTATGATCTACTGTTAAACCTAAAGCAACCAAGCTATTGAAGTTAAAGCAGAAAGTA 5557
QY 5702 TAGCAAAATGCTTTGTTAAAAG-----CAAAATCTCATTCATCTTA 5740
DB 5557 CCAAAAGAAATCTATTGATCAAGAGACCAAGTTAACTGTAAGAAAACCTGAAGCAATTAG 5617
QY 5741 TTTAAACAAATTGGAATTTGAGAGCTGGAAGTTGAGAGCTGAGAGAGTGAAGTGAAGTTCG 5800
DB 5617 CAATGTTTAAACAAATTTACAGATCAAGCTTAACAAAGGTAATTACTGTGCAACCAACTG 5677
QY 5801 TAGCAGTGAATAAGATTGTAATTAATACATAGCAGAAATTAATCATGCAAAAATCACTG 5860
DB 5678 CTGAAGTTAAAAAGGAAAGCTCAAGAGCTTGAACATTTGATTAACATTTCAATGAGACT 5737
QY 5861 CGAAGGAAATGCGAGATTATTACAGAGCTGATCGGTAATTTGCTTAATTTATGCAAGAA 5920
DB 5737 CAACAGAAAAACAAAAGCTATCGAAGAAATTAGAAAACCTGACATGACCGAATTTGAAGCAG 5797
QY 5921 CAGTGTCTGAGAGTCCCGTGCAGCAATGGAAGCTCAACCAAGTGAATGAATGAATTAACG 5980
DB 5797 GTGTAATGTCACGCTGATGTCACATGAAAGAAAAAGAGCGTTTACGAATGCTTTAG 5857
QY 5981 GATCTCAAAAAGCATATGTTAAAGATTTCTACAGTGAATGCTTAAGAAAGAAACAGATATT 6040
DB 5857 AAGACATTTTATCAAAAAGCACTGAAGATATTTCTGATCAAACTACAAATGCAAGAAATG 5917

```

```

QY 6041 ATATTACTACTCAAGGCAAGTAGATTAAGTGTAGATTAAGATTTCAAAAATCTTAATA 6100
DB 5918 CTACTGTCAAAAATAGTGCCTTGAACAACTTAAGACACAAAGATTAATCTGAAATTA 5977
QY 6101 TTAACGAAGACTTATACAAAAGAAATTAAGTAAATTAAGAAAGATTTGTTACCAATA 6160
DB 5978 AGAAAAATGCTTTGGAAGCAATCAGAGAAAGTGTAAACAAATTAAGAAATTAATAAA 6037
QY 6161 GTTCAAGCTACTCATACTT 6178
DB 6037 ATGAGATGATGATGAT 6055

RESULT 19
ACF73459
ID ACF73459 standard; DNA; 7434 BP.
XX
XX ACF73459;
XX
XX 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus DNA #1139.
XX
XX DE
XX AC
XX AC
XX DT
XX 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus DNA #1139.
XX
XX DE
XX AC
XX AC
XX DT
XX 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus.
XX
XX OS
XX PN
XX WO200294868-A2.
XX
XX 28-NOV-2002.
XX
XX 27-MAR-2002; 2002WO-1B002637.
XX
XX 27-MAR-2001; 2001GB-00007661.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Masignani V, Mora M, Scarselli M;
XX
XX WPI; 2003-120786/11.
XX
XX P-PSDB; ABW71899.
XX
XX DR
XX PT
XX PT
XX PT
XX PT
XX PS
XX
XX Claim 6; SEQ ID NO 2277; 49bp; English.
XX
XX The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is
XX useful as a pharmaceutical, particularly as a vaccine for treating or
XX preventing infection due to Staphylococcus bacteria, specifically an
XX infection caused by S. aureus. The composition is particularly useful for
XX treating or preventing sepsis in a patient. The composition can also be
XX used for diagnostics. The protein is also used in an assay for enzymatic
XX studies and as a target for antibiotics. This sequence represents one of
XX the novel S. aureus genes of the invention
XX
XX Sequence 7434 BP; 3195 A; 1326 C; 1300 G; 1613 T; 0 U; 0 Other;
XX
XX Query Match 0.8%; Score 79.2; DB 8; Length 7434;
XX Best Local Similarity 41.3%; Pred. No. 4,4e-05;
XX Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;
QY 3782 TTCAAGTAAACAGTGTGTTGTTAAATATGATATTCGAATGCAATTAATTTA 3841
DB 3659 TTAATTAATAAGCACTGCACTAATGAATACAGCAATTTTAATTAACAAATTTGCAAG 3718
QY 3842 AAGTAATGCTTTGAGTGAAGCAACTCAAGTACAGCAGACAGCTTTGGAAGCATTTA 3901
DB 3719 AGATTCAAGCTACGCGCAGATGCAACAGATGAAGAAACAAACGACGCTGATGCGAAGCAA 3778

```

Oy	3902	AAGAAAGTGAAGCAAGAAAAAGTTATCTATTGGAACTTGTGCTTATCACTTAG	3961
Db	3779	ATTACTGAAAAATGTGTAAGCAATCAAGCCATTTTCAGACGCACTTACCGCAAGTTG	3838
Oy	3962	TGAACAAATGAAGTTTCGCAAAATCGAATAATATACATGACGAGAAATCGAAAGCC	4021
Db	3839	ATGAAGCTTAAAGCAAAATGCAGAACGAGATTAATGCGTTAACCCAAAAGTTGTGAAG	3898
Oy	4022	AAAAAATGATGTTGATGTCACTGCTTATCAAGCGGACACCCAGTAGACAGAGCTTAA	4081
Db	3889	AACAAAGCGCTTAAGATTAATTTGATCAATTCAGACGCAACGCAAAATGTTATCATAT	3958
Oy	4082	ATTTCAGAGCTGGAAGTCAATGGAACCTGAGGCGCTACTGTGACTGTGCGCAATTAA	4141
Db	3959	ATGATTCAGAACGCTATCAACACAGAAAAAGAACGAGCTATTCAACAAATTGACACAGAC	4018
Oy	4142	ACAAACAAAGTAAATGCTCTTATTTAGTGTGGAAATATCTAACGTTAATCGAGCGACG	4201
Db	4019	TTACAGAGCGCAAAAAATTAATTTACAGTGCACATGATGAATGAAGGTGATGATCAGCGA	4078
Oy	4202	CAAAAGCTCTTTTAGCAACCACTCAAGTACTGCTGACAGTGCAGCGGAGCGCAATTA	4261
Db	4079	AAGAGCGCTGGAAGAAATTTCAATTCAAGACGCAACGCAACGCGTTTAATCAATG	4138
Oy	4262	GTTCTGAGCGGGAATTAGAAATTTATCAAGGCGGCTGTTCTGTCAATAGATTGACATG	4321
Db	4139	CTAAAAATGATGTTGATCAGCTGAGCACTCAAAATCAAGCAATTTGATTAACACTG	4198
Oy	4322	ACGTGGAAGCTAGCGTTGATTAATTTCTCATCGAAGAGCTAATGAATCAATGTCATTG	4381
Db	4199	GTGCTACCACTGGAAGAGAAAAATCAGCAAAAAAGTTTAGTTTAAAGCTTAAGAAAAAG	4258
Oy	4382	CCAAAGATGTCAAGGAAGTCTGATCTAGCAAAAGAAATACAGGCTTTACTAAATGAA	4441
Db	4259	CGTATCAAGATATCTTAATGACAAACACTTAATGATTTAGCCAAATTAAGATCAAG	4318
Oy	4442	AAGATAAAAAATATTTAGAGATCGTGTATTTAATACGACTGGAATGTTATTATACGA	4501
Db	4319	CAGTTGCTGATATTTCAAGGTAATTACTGAGATTAACAATAAAGATGTTG-----CGA	4372
Oy	4502	AGGAACAACCTGAAAAAGCAAAAGAAAAAGAGGCGGTATTGTAATGTCGCTTAT	4561
Db	4373	AAGATGAATTTAGCAACAAACGAAACGAAACGAAAGCGCTTATTGCAACAACTGCAGATG	4432
Oy	4562	CGGTGCTGGAACGGAATAATCCGCTGAGAGTAGCATTTGACGTCAATACTGTAAAA	4621
Db	4433	CGACTACTGGAAGAAAAAGAACAGAGAAATCAACAAGTAGACGCAAAATTAACACAGATA	4492
Oy	4622	ATTAATTTAAAGCAGAAATTGATGGAAGCATTAAGAAAGCGAGAGGATTAATTTGATG	4681
Db	4493	ATCAAAAAATTTGAAAAATGCAACATCTAATCGATGATGTAAACACTGCACAAAGATAATGCA	4552
Oy	4682	CGAAACATGTAAATGTGAGGCAAAATCATCTACTGTTGTTGTAATGCGGCTTCTGGAC	4741
Db	4553	TTCAAGCAATTAGCCCAATTCAGGATCAACAGATGTTAAAAAGAAATGCAAGAGC-GGAA	4611
Oy	4742	TTGCTATATAGCAAAAGATGCTTTTCAGGAATGGGATCTGGAGCATGGCAAGACTTATCA	4801
Db	4612	TTGCTAACTGAAAAATGCAAAATPAAATACTGAATACTTAATTAATGAAACG-----	4665
Oy	4802	ATGACACGATTGCAAAAGGTGATAAAGAAAGAAATTTCTGCTGATTTCCCTTAAATGTGAAG	4861
Db	4666	-----ACTTAATGAAGAAAAAGGTAAAGATTTTGAACCAAGTTTAGACAGCATATTAAGAG	4720
Oy	4862	CAAAATTAATTCATTTCTGCGGTGAATGTTCGCGGAACCATTCGCGGTTCTTTCTACGG	4921
Db	4721	GTTTAAATTAATTAATTAATGACAGAACTACTACAGGGATGTAACATACGTAAAGATACAG	4780
Oy	4922	CGGTAGGAGCTGCTTTTGCGAAATTAATCTTCAATATAAAACTCTGCTTTGATTAACAG	4981
Db	4781	CAGTACAAAAAGTTTCAACAACTTCTGCAAAATCTGTTTAGAAAAACACGACAGGTAAAAAG	4840

QY	4962	GAAGAGAGTAAATCCCTTTAGTGGAAAGAA	TACAAAGTCAATGTCAGGCTTGAATG	5041
Db	4841	AATTAGATCAAGCGTCAGCTGATAGAAAAC	ACAAATAGAACAAACACCAATCATCTAC	4900
QY	5042	ATTCTCATATATACAAACGTTTCTCTGAG	CGCTGCAAGATTAATACAGGCTGAAATCG	5101
Db	4901	AACAGAAATTTATATGATGCAAAACAAAG	ATTGATCTGAAATTTAACTCAAGCAAAACAA	4960
QY	5102	GAGGAATGATATCTGTCAATCGTGCTTCT	GATGAAACGAAAGCTTTAGTGAATTCGTG	5161
Db	4961	ATGTCGATCAATCATCAACAAATGAAAT	ATATGTTGTAATATGCAAGTTAAAGAAAG	5020
QY	5162	AGTTGGAAGGATGAAGTTCTTCAATGTA	TGCAAAAGATCAAAAACCAATTAATACAA	5221
Db	5021	AAATTTAATGCAAGTTTAAACATTTAGTA	GTACAAAAAGATGCTTTAGCTTAAATTAAG	5080
QY	5222	TTGCGCGAAATATGCAATGAGAGAAAG	CGCGCTGGAGTTGGAGCAACAGTTGCTCAT	5281
Db	5081	ATGCTATATATGCTTAAGGTAAACGAA	AGCGGATTAAGCTTAACGCTTCGACTTCA	5140
QY	5282	ATATTGGAAAACAATCAGTTATTAAGCT	ATTTGTAATAAAAAACAGTAAATTAACA	5341
Db	5141	TTGCTGAAGCGAAACAAAACCTTGCTG	AAATTTAAACAACTCGCGATCAAAATGTT	5200
QY	5342	ATCAAGATAGAAAAATATCAATGTA	CTGACGAAAGATTATCTATGACCAATCT	5401
Db	5201	AAGCTACTCTTCAAGATGACATTAAG	ATTCAAATTCATTAAGACTTGATATTTTA	5260
QY	5402	CAGTCGAGTTGGAGGAGCAAAAGAG	CGCTGTGCAAGAGGCTTTCGCAAGTCT	5461
Db	5261	ATTACACATTTCCACACAGTAAAAAG	AATAGCTACACAGATTTATATGCTTATG	5320
QY	5462	TGAATTAAGACAGTTTCTTCATGCT	GTGATCAATCTAATTTAGACAAAGATT	5521
Db	5321	ATCGAAGAAATATATATTTCACTG	ACACTAATGCAACAAAGATGAAAGCA	5380
QY	5522	AAAAATATGAAATTAAGAAAAAGCA	AAATGTTAATGTTCTAGCTGAAATAC	5581
Db	5381	CAATTTAAGCAAGTTGACCAAAATG	---TTCAATCTGATTAAGAAAGCAT	5437
QY	5582	TGATCACAATAGCGACAGTGC	TTTCCGAGCAATGACAAAGCTGAGT	5641
Db	5438	TGGATATATGAGCGTTGATGATG	ATTTACACAAAGTTAAAGCAATTTG	5497
QY	5642	TAGCAGTTAATAAATTAACAAAAAT	ATCTTCGACATATAAAAATAGTCTCA	5701
Db	5498	AAGTAGATGCTACTGTTAACTTAAG	CGAACCAAGCATTTGAAGTTAAAGCA	5557
QY	5702	TACGAATGCTTTGGTAAAAAG	-----CAATCTCATTCATCTA	5740
Db	5558	CGAAAGAAATCATGATCAAGTAG	CACAGTTAACTGTCGAAGAAAAAACT	5617
QY	5741	TTAAAACAATTGGAATTTGGAGCT	GTGAGCTGGAAGGCTTGGAGTGA	5800
Db	5618	CAATGATTTAAACAAATTTACAGAT	CAAGGTAAACAAGGATTTACGATGCA	5677
QY	5801	TAGCAGTGAATTAAGTTGTAATAT	AATACATAGACAAATTAATCACTG	5860
Db	5678	CTGAAGTTGAAAAAGCGAAAGCT	CAAGGACCTTGAAGCACTTTGAATTC	5737
QY	5861	CGAAGGAAATGTGCGAGTTATT	CAAGAGCTGATGCGGTAATTTGTAAT	5920
Db	5738	CAACAGAAAAACAAAAGCTAT	TCGAAGAAATTAGAAATGCACTAG	5797
QY	5921	CAGTGTCTGAGGTGCGCTGCGAG	CAATAGGAGCTCAACAGTGTGAATTA	5980
Db	5798	GTGTTAAATGTCAACGCTGAT	CTCAACTGAAAGAAAAAGACGTT	5857
QY	5981	GATCTACAAAAGCATATGTA	AAAGTTCTCAAGATGCTTAAGAAAG	6040
Db	5858	AAAGACATTTTATCAAAAGCA	CTGAGATATTTTGTGATCAACAT	5917
QY	6041	ATATTACTACTCAAGGCGC	AGTAGATTAAGGTGATTAAGGTAT	6100

DB 5918 CTACTGTCAAAAATGTCGGCTTGAACAACCTTAAGCACAACGATTTATCTCTGAAGTTA 5977
QY 6101 TTAACGAAGACTTATCAGAAAAAGAAAAATTAATTAATAAAGATTTGTTACCAATA 6160
DB 5978 AGAAAAATGCTTTGGAACCAATCAGAGAGTGTTAACAGCAAAATAGAAATTAATTTAAA 6037
QY 6161 GTTCAGCTACTCACTT 6178
DB 6038 ATGCAGATGCAGATGCAT 6055
RESULT 20
AAS55232
ID AAS55232 standard; DNA; 7437 BP.
XX AAS5232;
AC AAS5232;
XX 13-FEB-2002 (first entry)
DT 13-FEB-2002 (first entry)
XX Staphylococcus aureus DNA for cellular proliferation protein #1544.
DE Staphylococcus aureus DNA for cellular proliferation protein #1544.
XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KM antibacterial; drug design.
XX Staphylococcus aureus.
OS Staphylococcus aureus.
XX WO200170955-A2.
PN WO200170955-A2.
XX 27-SEP-2001.
PD 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009180.
PF 21-MAR-2001; 2000US-0191078P.
XX 21-MAR-2001; 2000US-0206848P.
PR 23-MAY-2000; 2000US-0207727P.
XX 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.
PA (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Tirawick JD, Carr GJ,
PI Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
DR P-PSDB; AAU37373.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX Claim 27; SEQ ID NO 8869; 511pp; English.
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes
XX themselves and the encoded proteins. The prokaryotes used are Escherichia
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence encodes an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX

SEQ Sequence 7437 BP; 3199 A; 1326 C; 1298 G; 1614 T; 0 U; 0 Other;
Query Match 0.8%; Score 79.2; DB 4; Length 7437;
Beet Local Similarity 41.3%; Pred. No. 4,4e-05;
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;
QY 3782 TTCAAGTAAACAGAGCTTGGTTAAATATAGATATTCGAATCCAAATTAATTTA 3841
DB 3659 TTAATTAATAAAGCACTGCACGTATATGAAATCAACGCAATTTTAATTAACAAATTTGCAAG 3718
QY 3842 AAGTAAATGCTTTAGTGAAGAACTCAAGTAGACAGAGAGAGGTTTGAAGCAGTTA 3901
DB 3719 AGATTCAAGCTACGCGACAGATCCAAACAGATGAAGAAAAACAGAGCTGATGCTGAAGCA 3778
QY 3902 AAGAAAGTGAGACAGAGAAAAAGTTATCTATTGGGAAGCTTCTCTATCAACTTAG 3961
DB 3779 ATACTGAAAAATGTAAAGCAATCAAGCAATTCAGAGCAACTAATACGCAAGTTG 3838
QY 3962 TGAACAAATGAAGTTCTGCAAAATCAGAAATTAATATACAGTACAGAGAGATCTGAAGCC 4021
DB 3839 ATGAAGCTAAAGCAATGACAGAGAGAGATTAATGCGTTAACACCAAAAGTTGAGAA 3898
QY 4022 AAAAAATGATGTTGATGTCATGCTTATCAAGCGGACCAACCAAGTGAAGAGCTTTAA 4081
DB 3899 AACAGCGGCTTAAGATGAATGATCAATTATCAAGAACCAACCAATGATTTATCAATA 3958
QY 4082 ATTTACAGCTGGAAGTCAAAATGGAAGTGAAGGAGCTTACTGTGCTTTGCGCAATTAA 4141
DB 3959 ATGATCAGAGCGTCAACAGAGAAAAAGAGAGCTTATTCACAAATTAGACAGAGAG 4018
QY 4142 ACAACAAATGAATGCTTCTTATGTTGGGAGATTAATACGTTAATGAGCGAGC 4201
DB 4019 TTACAGAGCGAAAAATTAATTAATGAGCTGCACTGATGATATGTTAGATCAGGCA 4078
QY 4202 CAAAAGCTCTTTTGAACAACCTCAAGTACTGCTGACAGAGAGAGAGCAATTA 4261
DB 4079 AAGAGCGTGAAGAAATTCATTCAAAGCAGCAACAGAGAGAGAGAGAGAGAGAG 4138
QY 4262 GTTCTGAGCGGAGTTAGAAATTAATCAAGGAGCTTTCTGCAATTAAGATTGACATG 4321
DB 4139 CTAAAAATGATGTTGATCAAGCTGACACTCAAAATCAAGCAATTTGATTAACACAG 4198
QY 4322 ACGTGAAGCTAGCGTTGATTAATCTTCATGAGAGAGCTTAATGAATCAATGTCATTG 4381
DB 4199 GTGCTACAACTGAAGAAAAATGACAGAAAAAGATTGTTTAAAGCTTAAGAAAAAG 4258
QY 4382 CCAAGATGTCAAAGAGATCTGATCTAGCAAAAGAAATATCAGGCTTTACTTAATGGA 4441
DB 4259 GGTATCAAGATATCTTAATATGACAAACACTAATGATGTTAGCAAAATTAAAGATCAAG 4318
QY 4442 AAGATTAATAATATTAGAAAGTCGTGATTAAATGACATGGAATGTTATTACGA 4501
DB 4319 CAGTTGCTGATATTCAGAGTATTTACTGCAATACAAATTAAGATGTTG-----CGA 4372
QY 4502 AGGAACAATGAAAGAAAGAAAGAAAGAAAGAGCGGCTTGTAAATGCTGTTAT 4561
DB 4373 AAGATGAATTTACCAACAAAGCAAGAAAGAAAGCGTTTATGACAACTGCGAGATG 4432
QY 4562 CGTTGCTGAGACGATTAATCCGCTGAGAGATGATTTTACAGTCAATACGTTTAA 4621
DB 4433 CCACTACTGAGAAAAAGAAAGAAAGCAATCAACAGTATGACCAATTAACACAGATGA 4492
QY 4622 ATTAATTTAAGCAGAAATTTAGTGAAGCAATTAAGAGCCGAGAGATTAATTTATG 4681
DB 4493 ATCAAAATTTAATAATGACAGATCAATGATGATTAACCTGCAAAAGTAAATGCA 4552
QY 4682 CCAACATGTAATGAGAGCAAAATCACTACTGTTGTTGATGAGGCTTGTGAGC 4741
DB 4553 TTCAAGCAATGACCAATTTCAAGATCAACAGATGTTAAACAGATGCAAGC-GGA 4611
QY 4742 TTGCTATCAGAAAGATGCTTTTTCAGATGAGATCTGAGACATGCAAGACTTATCA 4801
DB 4612 TTGCTATCAGAAATGCAAAATTAATACTGAATATCTTAATTAATGAGACT----- 4665

[illegible]

XX	RESULT 22
XX	AAA70259
ID	AAA70259 standard; DNA; 4677 BP.
XX	
AC	AAA70259;
XX	
DT	07-NOV-2000 (first entry)
XX	
DE	Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:392
XX	
KW	Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine; antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX	
OS	Plasmodium falciparum.
XX	
PN	WO200025728-A2.
XX	
PD	11-MAY-2000.
XX	
PF	05-NOV-1999; 99MO-US026796.
XX	
PR	05-NOV-1998; 98US-0107131P.
XX	
PA	(HOFF/) HOFFMAN S.
PA	(CARU/) CARUCCI D.
PA	(GARD/) GARDNER M.
PA	(VENT/) VENTER J C.
XX	
PI	Hoffman S, Carucci D, Gardner M, Venter JC;
XX	
DR	WPI; 2000-365347/31.
XX	
PT	Proteins encoded by chromosome 2 of the human malarial parasite,

PT Plasmodium falciparum, useful as antimalarial vaccines and in the diagnosis of P.falciparum infection.

PS Disclosure; Page 565-566; 577pp; English.

The present invention describes proteins and their fragments (I) encoded by chromosome 2 of the human malarial parasite, *Plasmodium falciparum*. Also described are: (1) nucleotide sequences (II) encoding (I); and (2) vaccines against *P. falciparum* infection comprising (I) or (II). (I) and (II) are useful for the development of vaccines against *P. falciparum* infection. (I) and polyclonal antisera or a monoclonal antibody raised to immunogens comprising the sequences of (I), are useful in the detection of infection with *P. falciparum*. Furthermore, (I) (especially when they are rifins or secreted or membrane proteins) can aid the identification of drugs to treat or prevent *P. falciparum* infection, or they can be used to identify drug resistance in *P. falciparum*. Sequencing of the *Plasmodium* chromosome 2 and the subsequent identification of proteins encoded by it will help to expand our understanding of parasite biology, a process hampered by the complexity of the parasitic lifecycle, and provide new targets for vaccine and drug development. Parasite resistance to drugs and mosquito resistance to insecticides have led to a resurgence of malaria in many parts of the world, and there is a pressing need for vaccines and new drugs. AA70078 to AA70287 and AB18144 to AB18352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification

Sequence 4677 BP; 2106 A; 402 C; 966 G; 1203 T; 0 U; 0 Other;

Query Match	0.88;	Score 76.4;	DB 3;	Length 4677;
-------------	-------	-------------	-------	--------------

Matches 1538; Conservative 0; Mismatches 2161; Indels 45; Gaps 12;

5496 TCATATTGACAAAGATTTAGAGGAGAAATAATGGAAATAAGGAAAGCCAATGTTAA 5555

Db 375 TAATAAAATTTAGAGAAGCCGAGATATTAAGGAAATATCTTATTAGTAATATAGA 434

5556 TGTTC TAGCTGA AATACGAGTCAAGTGGTCACAAATGCGACAGTGTTC CGGAGCAAG 5615

Db 435 AGAACCAAGAAATATTATTGACCAATTATTAAATAATTTGACAAATTGAGAAA 494

5616 TGGACAAGCTGCAGTAGGAGCTGGAGTAGCAGTTAATAAATTACACAAATACTTCTGC 5679

Db 495 ACAGAAAGTGTATCAGAAATGTACAAGTCAGTGATGAACCTTTTAAATGAATTATTA 554

5676 ACATATAAAATAGTACTCAAATGTACGAATGCTTGGTAAAGCAATCTCATTC 5735

Db 555 TAGGTAGATGTTAATGGAGACTAAAGAAATATTTTGGAGGAAAGTCAAGTTAATGA 614

5736 ATCTATTAAACAATTTGGAA TTGGAGCTGGAGCTGGAGGAGCTGGAGTGCACAGG 5799

Db 615 CGATATTTTAAATAGTTTAAAGTGTTCACACAGACACACATGTTGAAGA 674

5796 TTCTGTAGCAGTGAATAAGATTGTAATAATACGATAGCAGAACTTAATCATGCAAAAT 5855

Db 675 AAAAGTTGAGAAAGTGTAGAGAAATGACGAGAAAGTGTAGAGAAATGTAGAGA 734

5856 CACTGCGAAGGAATGTCGAGTTATTACAGAGTC--TGATCCGTAATTGCTAATTA 5912

Db 735 AAATGTAGAGAAATGACGACGAAGTGTAGCCTCAAGTGTGAAGAAATAGCTTC 794

5913 TGCAGGAACAGTGTCTGGAGTGGCCCGTGCAAGCAATAGGAGCCTCAACCACTGTGAATGA 5977

Db 795 AAGTGTGATGAAGTATAGATTCAAGTATTGAAGAAATGTAGCTCCAACGTGGAAGA 854

5973 AATTACAGGATCTACAAAGCATATGTAAGATTCTACAGTGATTGCTAAGAGAACA 6032

Db 855 AATCGTAGCTCCAAGTGTGAGAAATTGTAGCTCCAAGTGTGTAGAAAGTGTGGCTCC 914

6033 AGATGATTATTTACTACTCAAGGCAAGTAGATAAAGTGTAGATAAAGTATTCAAAA 6097

Db 915 AAGTGTGAGAAAGTGTAGAGAAATGTTGAGAAAGTGTAGCTGAAATGTTGAGA 974

QY 6093 TCTTAATATTAAGAAAGCTTATCACAAAAAGAAAAAATAGTAAATAAAAAGATTGT 6152
Db 975 ----AAGGTAGCTGAATAATGTTGAAGAAAGGTAGCTGAATAATGTTGAAGAAAGGTAG 1030
QY 6153 TACCAATAGTTGAGCTACTCAATCTTTAAATCTTTATTTGGCAATTCGCGCTGGTTCCAG 6212
Db 1031 CTGAAAATGTTGAAGAAAGGTAGTGAAGAAATGTTGA--AGAAAGTGAAGTGAAGAAATGT 1088
QY 6213 ACAAGCCGAGTGGCAGAGACTGTTTAATATCAACAAGGTTTATGAGAAAACAGAGCTCT 6272
Db 1089 TGAAGAAATCGTAGCTCCAACTGTTGAAGAAAGTGAAGTGAAGTGAAGAAATGT 1148
QY 6273 TGTAGAAAATCTATATTAATTAATGCAAAACATTATCTGTAAATCAGAGATTAACAGCA 6332
Db 1149 AGCTCAAGTGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1208
QY 6333 TTCAATCGAGTGAAGTGTCTGT 6392
Db 1209 TGAAGAAAGTGAAGTGAAGAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGT 1268
QY 6393 TGATACCAATATTAATTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 6449
Db 1289 AGCTCAAGTGTGAAGAAATGTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1328
QY 6450 TGATGAAGTGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 6509
Db 1329 TGAAGAAATGTTGAGCTCCAAAGTGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1388
QY 6510 TTTTGAAGTGGAGTGGCAGAGCCGGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6569
Db 1389 AGTTCAACTGTTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1448
QY 6570 AAATCAATTTTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6629
Db 1449 TGAAGAAATCGTAGCTCCAAAGTGTGAAGAAATCGTAGCTCCAAAGTGTGAAGAAATGT 1508
QY 6630 AGCTGAGATTAACAGAAACGTTATAGTTGTTGCAATTGGAATGCGCGAGTGAAGT 6689
Db 1509 AGCTCAAGTGTGAAGAAATCGTAGCTCCAAAGTGTGAAGAAATCGTAGCTCCAAAGTGT 1568
QY 6690 GCGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6749
Db 1569 TGAAGAAATCGTAGCTCCAAAGTGTGAAGAAATCGTAGCTCCAAAGTGTGAAGAAATGT 1628
QY 6750 GAGAGCAGAGTGAAGAAATCTTAATATTAATGACTGCAAAACAGTAGATG-----TAAT 6803
Db 1629 AGCTCAAGTGTGAAGAAATCGTAGCTCCAAAGTGTGAAGAAATCGTAGCTCCAAAGTGT 1688
QY 6804 AGCAGAAATGAGATTAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6863
Db 1689 TGAAGAAATGTTGAGCTCCAAAGTGTGAAGAAATCGTAGCTCCAAAGTGTGAAGAAATGT 1748
QY 6864 AGCCGAGATTAATGAGTGAAGTGTCTGTCAATTAATTCGCAATTAAGTGAAGAAACAGATAT 6923
Db 1749 TGCCTGAAAACGTTGCAAAATTAATCAAGCAATCTTTTAAGTAAATTAATGAGTGAAT 1808
QY 6924 CGATCATAGTACTTACACTCTCTTCTACTGATGTAATGTAATGTAATGTAATGTAATGTAAT 6983
Db 1809 CGAAATCAAGAAATTAAGACAGATTAATTAATGATTAAGAAAGTGAAGAAATGTT 1868
QY 6984 GAAATTCCTTGAACAGCGGTGAAGAGCCGAGGTCTTGAAGCAAGTGAAGTGAAGTGAAGTGAAG 7043
Db 1869 AGTACCAATATTAAGAAAGTGAAGAAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1928
QY 7044 TGTTAACCTATTAATAGTCTGTGTATAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7103
Db 1929 TAACTATTAAGAGGATACAAAGAAATATTAATTAATGATTAATGAGTGAAGTGAAGTGAAG 1988
QY 7104 CGTACGAGAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 7163
Db 1989 AGAAAGATCTCCAGCAAAATGTAATTAAGTGCCTTTGAAGAAATCAAGTGAAGTGAAGTGAAG 2048
QY 7164 AAATGACAGAAATCGAGAGAGCAGCAATCGAGCAATGTCTTGTAAATTAATTTTGAAG 7223

Db 2049 AAAGAAAGATTAATAGATGTAATGTAAGAAAGTAAAGAAAGAGGTGCTACCACTTTAAT 2108
QY 7224 AGCTGTAGAAATGAAAAAATTTCTGAAGAAAGAAAGAAACAGAGTTTAAAACTTTAGA 7283
Db 2109 AGAAATGTTGAACAGCAGAGAAAGAGAGCGCAAGTCAATTAACGAAATATTTGAAAAA 2168
QY 7284 CGAAGTTAACAAAGAACAGATTAAGAAATGATGCTGCAAGAAATTTTCAATTC 7343
Db 2169 TTTGAAGAAATATCAGTGAAGAAATGTAATGAAATGTTGCAAGAAATTTTGAAGAAATTTAA 2228
QY 7344 AGCAGTATTTCTACAGAAATCTCTGTGAAGAAAGCGATTAAGAGATTAAGTGAAGTGAAG 7403
Db 2229 CGAAATCTGATTAATAGTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2288
QY 7404 AGGAATTAAGCCATTTGGAAGACTTCTGATTAATTAAGTGAAGAAATGTAATTAACAC 7463
Db 2289 AAGTTTGAAGAAACAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2348
QY 7464 AGAGACAAAGATTAATATCACTTCTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7523
Db 2349 AGGAATACAGAAATTTT--ATTACAGGTATGTTTGAAGTATGAAGAACAGATATGT 2405
QY 7524 ATCAGAAACAGTGCAGTTACAAATATTAAGAAATTCGAGTGAAGTGAAGTGAAGTGAAGTGAAG 7583
Db 2406 AATCCAAATCAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2465
QY 7584 TTTTGTGAAGCAGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7643
Db 2466 TATGAAATATGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2521
QY 7644 AACGACATCAAGTCTGTGAGAGCAATGGGAATGAGAGTCTGTGAGAAATTAAT 7703
Db 2522 --AAGTGTGAAGAACTGTACTGAATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2579
QY 7704 TTTTAATGAAGATCAATATCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7763
Db 2580 TGTTCCTGTATGAAGATCAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2639
QY 7764 TGTATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7823
Db 2640 AATGTTTTTAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2699
QY 7824 GGTAGTGCAGGACCAATATCTCAAAAGCAAGAAATGAATGAATGAATGAATGAATGAATGAAT 7880
Db 2700 AATTAAGATGAACCGGTCTCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2759
QY 7881 AATGAAGAGTATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7940
Db 2760 AATGAAGAAATTAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2819
QY 7941 AAGGAAATCAATGTCAGAGTGAAGAAAGAAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8000
Db 2820 GATGATGCAAGTGAAGAAATCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2879
QY 8001 TTTCTGAGAGCAGTGAAGAGGAGGAGGAAATTAATTCGAGCAAAAGATGCGGAGGCTC 8060
Db 2880 TATTAAGATTAAGAAAGATGTTTCACTAGTGTGAAGAGTTCAGAGCAATGATAT 2939
QY 8061 TTAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8120
Db 2940 GATGAAAGAGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2996
QY 8121 GGAAGCAACATTAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8180
Db 2997 GGAATGCTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3056
QY 8181 GGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8240
Db 3057 TGAAGTGAAGCAATTTTAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3116
QY 8241 TGAAGAGAAATTTGTTCAAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8300


```

Db      467 CCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 526
Qy      564 CATTACGGTAGAGAGAAATCAATGCTGTTGAGGACGCTGGTTATATGCGCGGATAT 623
Db      527 AATTAAAAAAAAAAAAAAAACAATTAATAAAAAAAAAAATTAATAA-AGATTAAAAAAAAAAAAAGT 585
Qy      624 TAGATTGAAGATGACTGCAATCTAAGACAGAAATTAAGATTTTAAATAATTTAGTCA 683
Db      586 AAAATACATTAATTTAAAAACAAAAAAGAAAAACNCGCATTAATAATTAATAATAA 645
Qy      664 TATTGTGATGCAATTAATTTCTGCTGACCGGAGATTAAAAAGCTACCAAGACAAATC 743
Db      646 CTATTAATAAACAATAAAGTGTCTTTCTTAAAAAATAATCTCCCTTAAAAACA 705
Qy      744 TGGAGATATTTATTTCTTTCAGCTCAGATGATCTCCTCAAAAAGC---TATGGGAAAA 800
Db      706 TAAAAAATAAATAAAGAAATTAACAAAGGCTCTAATCAACCCCAATATTAACAAAA 765
Qy      801 TTCAACTGTGGAAAGAGATAGAAATATGTAAAGAAATACCAAGCAAAATATTTGA 860
Db      766 TCTTAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 825
Qy      861 ATCTGATGCTGATTTGGAAGCAGATGGAATATTAATAATTTAGTGGAAAGCTACAAATG 920
Db      826 AACTTAAAAAATGAGGGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 885
Qy      921 GAGATTTTAAAGAAAGAGGGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 980
Db      886 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 945
Qy      981 TGTGGAAGCTTCCTGAGAGTAATTAAGAAAGTCAATAGAAAGAAATGTTGACATTAC 1040
Db      946 TAAAAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1005
Qy      1041 AGCTGAGCAAGA 1054
Db      1006 AAAAAAAAAAAAAA 1019

RESULT 24
AAS51998
ID      AAS51998 standard, DNA; 7104 BP.
XX      AC
XX      AAS51998;
XX      DT      13-FEB-2002 (first entry)
DE      Staphylococcus aureus DNA for cellular proliferation protein #415.
XX      KM      Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX      KM      antibacterial; drug design.
XX      OS      Staphylococcus aureus.
XX      PN      WO200170955-A2.
XX      PD      27-SEP-2001.
XX      PF      21-MAR-2001; 2001WO-US009180.
XX      PR      21-MAR-2000; 2000US-0191078P.
XX      PR      23-MAY-2000; 2000US-0206848P.
XX      PR      26-MAY-2000; 2000US-0207727P.
XX      PR      23-OCT-2000; 2000US-0242578P.
XX      PR      27-NOV-2000; 2000US-0253625P.
XX      PR      22-DEC-2000; 2000US-0257931P.
XX      PR      16-FEB-2001; 2001US-0269308P.
XX      PA      (ELIT-) ELITRA PHARM INC.
XX      PI      Haselebeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GT;
PI      Yamamoto RT, Xu HH;

```

```

XX      WIPI: 2001-611495/70.
DR      P-PSDB; AAU34139.
XX      PT      New polynucleotides for the identification and development of
XX      PT      antibiotics; comprise sequences of antisense nucleic acids.
XX      PS      Claim 27; SEQ ID NO 4580; 511bp; English.
XX      CC      The invention relates to antisense inhibitors of genes essential to
XX      CC      prokaryotic cellular proliferation, their use in identifying the genes,
XX      CC      their use in the discovery of novel antibiotics, the essential genes
XX      CC      themselves and the encoded proteins. The prokaryotes used are Escherichia
XX      CC      coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX      CC      Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX      CC      useful for the identification of potential new targets for antibiotic
XX      CC      development. The antisense nucleic acids can also be used to identify
XX      CC      proteins used in proliferation, to express these proteins, and to obtain
XX      CC      antibodies capable of binding to the expressed proteins. The proteins can
XX      CC      be used to screen compounds in rational drug discovery programmes. The
XX      CC      antisense nucleic acid sequence is also useful to screen for homologous
XX      CC      nucleic acids which are required for cell proliferation in a wide variety
XX      CC      of organisms. The present sequence encodes an essential prokaryotic
XX      CC      cellular proliferation protein. Note: The sequence data for this patent
XX      CC      did not form part of the printed specification, but was obtained in
XX      CC      electronic format directly from WIPO at
XX      CC      ftp.wipo.int/pub/published_pct_sequences
XX      SQ      Sequence 7104 BP; 3045 A; 1258 C; 1552 T; 0 U; 0 Other;

Query Match      0.8%; Score 74.8; DB 4; Length 7104;
Beet Local Similarity 41.6%; Pred. No. 0.00031;
Matches 992; Conservative 0; Mismatches 1367; Indels 27; Gaps 7;

Qy      7208 TAAATATTGGAACGAGCTAGAAATAGAAAAATTTGAAAGAAAAAGAAACAGAG 7267
Db      2702 TCAATGCAATTCAACACAAACAAAGTTAAACCTGACGCTGATCGGAATGAAAAACG 2761
Qy      7268 TTTTAAAAAAGTTTGAAGAGTTAAACAAAGAAAGAAATTAATAATGATGCTACGA 7327
Db      2762 CATTAATATACGTTAAACAAAGAAATTCAAAATAGCAATGCTTCACTACAGAAAGAAAC 2821
Qy      7328 AAAAAATCTTCAATCAGCAGAGTATTTCTACAGAAATGATCTTGTAAAAAGCGATAGAG 7387
Db      2822 AAGCTGATATACAGAAATTAATGATCTAATAAAGCAAGAAAGCAAAATCTGATGCTG 2881
Qy      7388 GAGATCTCAGGAGAAAGAAATTAAGCATTTGAGAAAGTCTTGATATTTTGGAAAAA 7447
Db      2882 CAAATTAACAACAGTGAAGTAAACAGCTAAAGCAATGATGCTGCAATTAATCAAG 2941
Qy      7448 ATGTAGATATTACAAACAGAGACAAGAAATATATCACTTCTAGTGTGTTGGAACTG 7507
Db      2942 TACAAGCGCAACAATTAAGAAATCGATGCTTAAGCGAAATGCGTCAAAAACCAAGT 3001
Qy      7508 CAGGTCTTCTCCGCAATCAGAAACAGTGCAGTTACAAATATTTAAAGAAATTCGGAG 7567
Db      3002 AAGCTAAACT---GCAATTAACAATGAATGATGCACTACGAAAGAACAAAGCTG 3058
Qy      7568 TTACTGTGAAATTTCTTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGATCGATATTA 7627
Db      3059 CAAAAGATAAAGTTGATCAACAGATGATCTGCAAAAGCTGATATATGATATCTGCGAG 3118
Qy      7628 CAGGAATGTTGCTTTAAACAGCATATCAAGGTCTGTAGAGCAATTTGGGAATTAAGAGCTG 7687
Db      3119 CAAATCTGATGTGATGATATCAAAAATCTAATTAAGAGCTACATCGACCACTTACAC 3178
Qy      7688 CCTATGCAATTAATTTCTAATGGAAGATCAAAATATCTAGTATTAATAATTTCTAAGCTAT 7747
Db      3179 CTGATGCAAAATGTTAAACCAACAGCAAAACAAGCAATGCTGATTAAGTATCAACGCGAAG 3238
Qy      7748 TAGGAAAAATATGATGTTA---TTGTAAAAAGTAAATCGAAATTTGAGACCGGAAGCA 7804
Db      3239 AAACAGCAATTGATGCTAATTAACGATGCAACAGAAAGAAAGAAAGCAAGCTCGAAACAAAC 3298

```

QY 7805 AAGATTACCGTAGAGACGGTAGTCTCCGAGCCATTATCTCAAAAGCAAGATGAAA 7864
 Db 3329 AAGTTCAAACTGAAAAAACAACAGCTGATACAGCAATTGATGTCACATACAAATGAG 3358
 QY 7865 TGAATTCAGAGGTGAAATTTGAGAGAGATATTTTCAATGAAAGAAATAGATTAAGCC 7924
 Db 3359 AAGTTGAAGCGGCTAAAAATGCAAAATTTGCTAAAATTTGAGCAATTCAGCCAGCA 3418
 QY 7925 CTTCGAAAGAAATG--GAGAGAAATCAATGTCAAAAGTGAAGAAAAAGAAAAAGAGTGA 7981
 Db 3419 CAATTAAGATTAATCGAAACAAGCAATGTGACAAAGCGAATGAACTGTAACAGCA 3478
 QY 7982 CTGCTGAATCTCAAGAGCTTTGTAGAGAGCTAGCAGGCGGAGAGATTAATTCGAG 8041
 Db 3479 TCGCTCAACCGCAACATTAATCTGCTGAAGAAATTTGACGCGGCAATGCGAAT--GTAG 3535
 QY 8042 CAAGAATGCGGAGCTCTTAATTTGAAAGTTAGTACAAATCCGGAAGAGTATTTTTC 8101
 Db 3536 ATAACTGCTGTGACACAGCAAAATTAACACATTTGAAGCTCTAATAGTCAAAATGATGAG 3595
 QY 8102 ATGAGATTAATGTAATAGAGCAACATTAATGAAAGTAAACAGCATTTCTTAAG 8161
 Db 3596 ACCAAGCAAAACAATCTGCTGAGAGTAAATTTGATCAAGTAAACCAACAGTTAATTA 3655
 QY 8162 CAGTACAGAGTTCTGTATTGAGAGAGTTGAGTCAACAGCAGAGCTACTGCTGAG 8221
 Db 3656 AAGCAACAGCATTTACAGACGCAAAATTAATTAACAGTCACTGATGATTAATGAGT 3715
 QY 8222 GTAAATCTATGTTAGAGTTGAGAGAAATTTGTCAGAACTAATGATTAATGCA 8281
 Db 3716 TAGATACAGCGAAAGTGAAGTAAATTTGATTAATTAATTAATTAATTAATTAAT 3775
 QY 8282 TTTCTAAAGTAAAGTTGATGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 8341
 Db 3776 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8385
 QY 8342 GAAATGAGAGAGATTTGCGGAGCAGAGTGAATCTTTACAGCAAAAGTAAATCTG 8401
 Db 3836 ATAAATCAACTGTGCTACAACTGAAGAAAAATGCAAAAGTTTAAATTTTAAAG 3895
 QY 8402 AATCGTATGTTGTTAAGAAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 8461
 Db 3896 CTAAAGAAAAAGCTATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3955
 QY 8462 TTTTCAAGATCAATGCTCTTGTAAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 8521
 Db 3956 TTAAGATCAAGCAAGTTGCTGATTTCAAGATTTTCTGCAATTAATTAATTAATTAATTAAT 4015
 QY 8522 TAGCGTAGCCGCTGTGATGCAAGAAACAAACAAAGCATTTACGATCAACACAGT 8581
 Db 4016 TTGCGAAAGATGAATTAGCAACAAAGCAAGGAAACAAAGCACTTAATTTGCAAAATCG 4075
 QY 8582 TAACTTCTAACCTGTAATGAGAGAAAGTATCTCAATCTGCGAATAAGCTTTGGCTA 8641
 Db 4076 CAGATCGCTATGGAAGAAAAAGCAAGCAAAATCAAC-----AAGTGAATGCAAGAT 4129
 QY 8642 AAAATGAAATTTAGAAATGTAAGAACTGAGAGAGCTTAATGCTGAGCGGAAACAG 8701
 Db 4130 TAAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4189
 QY 8702 CAGCCGTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8761
 Db 4190 AAGATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4249
 QY 8762 TTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8821
 Db 4250 CAAAGACGGGATTTGCTAATGAAATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 4309
 QY 8822 GAACTAAAGAGCTCTTGTGCGAAAGATGTAATTTCTGTAAGAAATTAATTAATTAATTAATTAATTAAT 8881
 Db 4310 AGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4369

QY 8882 AAACAAATCATCATTTGAATTAAGCAGAAATTTGTTGAAACCGAAGTGAATGTAAG 8941
 Db 4370 GTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4429
 QY 8942 ATGCTTTGAATGAATTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9001
 Db 4430 CAGTACAAAAAGTTCAACAACTTCATGCAAAATCTGTTAGAGAGCAGCAGTAAATCTG 4489
 QY 9002 GTATTGGAATTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9061
 Db 4490 CTTTAGATCAAGCTGACCTGATTAAGAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4549
 QY 9062 GACATGCTATTGTAAGAACTACTGGAACAAAGATATCAAGATTTTCAAGAGCAAAAG 9121
 Db 4550 AACAGAAATTAACGA---TGCAAAACAGAAAGTTGATTAATTAATTAATTAATTAATTAATTAATTAAT 4606
 QY 9122 TAAATATTTCTTGAAAGAGAGCGCTGACAGCTGCAATGATTAATTAATTAATTAATTAATTAATTAAT 9181
 Db 4607 CAATATTAATGACCAATCTTCAACAGATGAATGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4666
 QY 9182 CCAATGAGATGATTTTAAATTTGCAAGAGATATGCAATCTTCTCAATTAATTAATTAATTAATTAATTAAT 9241
 Db 4667 CTAAATTAATGCAAGTTAAACATTTAGTGAATCAAAAGAGATCTTTAGCTTAAGATTG 4726
 QY 9242 AAAATTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9295
 Db 4727 AAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4786
 QY 9296 TTCAATGAGGAGCTGAGCAAGAGTGAAGAGCCAAAGCGACAGTTAGTAAAGATG 9355
 Db 4787 AAATGCAAGAGCGAAACAAATTTGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4846
 QY 9356 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9415
 Db 4847 ATCAAGCTAATCTTAAGAGTGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4906
 QY 9416 ATGTAATGCGGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9475
 Db 4907 ACGATTTACCAATTTCCACAGGTAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4966
 QY 9476 CGGATGCCAAAGTCAATGCTGAGCTGCTTGGCAATGCACTAATTTGAAATTAATTAATTAATTAATTAATTAAT 9535
 Db 4967 CAGATCAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5026
 QY 9536 TAAATTTAATTAATGCGATCGAGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9581
 Db 5027 AAGCAATTAAGCAAGTTGACCAAAATGTTCAAACTGCAATTAAGAAAG 5072

RESULT 25
 AAS54654
 ID AAS54654 Blandard; DNA; 7107 BP.
 XX
 AC AAS54654;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Staphylococcus aureus DNA for cellular proliferation protein #966.
 XX
 KW antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 KW antibacterial; drug design.
 OS Staphylococcus aureus.
 PN MO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JM, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX

DR MPI: 2001-611495/70.
XX P-PSDB; MAU36795.

PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.

PS Claim 27; SEQ ID NO 8291; S11pp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pcc_sequences

XX Sequence 7107 BP; 3047 A; 1258 C; 1249 G; 1553 T; 0 U; 0 Other;

Query Match 0.8%; Score 74.8; DB 4; Length 7107;
Best Local Similarity 41.6%; Pred. No. 0.00031;
Matches 992; Conservative 0; Mismatches 1367; Indels 27; Gaps 7;

QY	7208	TAAATATTTTGGAAACGCTGTGAAAGATGAAAAATTTCTGAGAAAAAGGACAGAG	7267
DB	2702	TCATAGCAATTCACAAACGAAAGAAAGTTAAACCTGACGCTGATACGGAAAGTAAAGAACG	2761
QY	7268	TTTTAAAACTTTAGACGAAAGTTAAACAAAGAACAAAGTAAAAAGTAAATGATGCTACGA	7327
DB	2762	CATATATATACAGTAAACAAAGAAATTCAAAAATGCAATGCTTCAACTACAGAAAGAAAC	2821
QY	7328	AAAAAATCTTCAATCAAGCAGATTTTCTACAGAAAGATCTTCTGTAAGGCGGATAG	7387
DB	2822	AGCTGCATATACAGATTAATGATTAATAAAGCAAGAACAAATCTTGATGCTG	2881
QY	7388	GAGATACCTCAGGAGAAAGAAATTTAAAGCCATGTGAAAGACTCTGATATTTATGGAAAA	7447
DB	2882	CAAAATCAACAGATGATGTAACAACGCTAAAGACATGCTATCTGCAATTAATCAAG	2941
QY	7448	ATGTAGATATTACACAGAGAACAGAAATTAATATCACTTCTACGTGTGTTGGAACTG	7507
DB	2942	TACAAAGCGGCAACAACTAAGAAATCGATGCTAAAGCGGAAATTCCTCAAAAAGCAAGG	3001
QY	7508	CAGGCTTGCTCCGATCAGGACAGTGGCAGTTTACAATATTTAAAGAAATTTCCGGAG	7567
DB	3002	AAAGTAAAGCT--GGAATTTGAAGCAATGATGATTCGATCTAGAAAGCAACAGCTG	3058
QY	7568	TTACTGTGAAATTTCTTTGTGAAAGCAGCTGAAAAAGTAAATGTATGATCGATATTA	7627
DB	3059	CAAAAGATTAAGTATGATCAAGCAGTATGATCTGCAAAAGCTGATATAGTATATGCTGAG	3118
QY	7628	CAGGAATGTTGCTTTACAGCATATCAAGGCTCTGTAGAGCACTGGGAATAGAGCTG	7687

DB	3119	CAAACTAGATGTAGATATATGCAAAAACCTACTAATGAAGCTACAAATCGACCATTTAC	3178
QY	7688	CCTATGCAATTAATTTCTAATGGAAGATCAAAATATCAGTATTAAAAATTTCTAAGCTAT	7747
DB	3179	CTGATGCAATGTTAAACCAACAGCGAAACAGCAATTTGCTGATTAAGTATACAGCGCAAG	3238
QY	7748	TAGAAAAAATATTGATGTTA---TTGTAAGAATTAATTCGAAATTTGAGACCGAAGCA	7804
DB	3239	AAACAGCAATTTGATGCTAATTAACGATGCAACAAGAAAGAAAGAACAGCTCGAAACAC	3298
QY	7805	AAAGATTACCGTGAAGAGCGGTAGCTCCGAGACCTATATCTCAAAAGCAAAAGATGAA	7864
DB	3299	AAAGTTCAAACTGAAAAAACAACAGCTGATACAGCAATTTGATGTGACATCAATCAATCG	3358
QY	7865	TGAATTCAGAGTTGAAATTTGAAAGAGTATTTTCAATGAAAGAAATAGATTAATACC	7924
DB	3359	AAAGTTAAAGCGGCTTAATTAATCAGAAATTTGCTAAATTTGAAAGCAATTTAGCCACAC	3418
QY	7925	CTTCAAAAGAAATTTG---GAAAGAAATCAATGTCAAAGTGGAAAAAGAAAAACAGAGTGA	7981
DB	3419	CAACTAAAGATTAATGCGAAACAAGCAATTTGCTAGAAAGGAAATGAACTAAACAGCA	3478
QY	7982	CTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTACAGGCGCAGAAATTTTCCGAG	8041
DB	3479	TGCTTCAAAACCAAGACATTAATCTGTGAAAGAAATTTGCGCGCAATTCGAAAT--GTAG	3535
QY	8042	CAAAAGATGCGGAAGCTTTATTTGAAAGTATGACAAATTCGGAGAAAGTATTTTC	8101
DB	3536	ATAATGCTGTACACACAGCAAAATTAACAACTTTGAAGCTCTAATATGTAATGATGTAG	3595
QY	8102	ATGCAATATATGATATGAAATGAGCAACATTAATTAATGAAATGACAGCTTTCTAAG	8161
DB	3596	ACCAAGCGAAAAACAACGTGTGAAGTATGATTAATGATCAAGTAAACCAACAGTATTA	3655
QY	8162	CAGTACAGGCTTCTGTATTTGGAAGAGTTGAGATGCCAAGCAGAAAGCTACTGCTGAG	8221
DB	3656	AAAGAACGACGATTACACAGCCGAAATAATATTTACAGCTGCAACTGATATATATGTG	3715
QY	8222	GTAATACTATGTAAGTATGAGAAAGAAATTTGTTGAGAAACAAATGATTTGAATGCA	8281
DB	3716	TAGATACGCAAGAGTGCAGATTAATTTGATTTCAAGATTAACAGTACACACAGCAGCG	3775
QY	8282	TTTCTAAAGTGAAGGTTTGGATGAGATTAAGTAACTGCTAAATCTTCTGATATGAG	8341
DB	3776	TTAATCAAAATGCTAAAAATATGATGATTCAGAGCTGGAACAATCAAAATCAAGCAATG	3835
QY	8342	GAAATGAGAGGAATTTGCCGAGCAGAGTGAATCTTCTACAGCACAAGTAACTAG	8401
DB	3836	ATAATACAACTGTGTCTACAACTGAAGAAAAATGCGCAAAAGATTTAGTTTAAAG	3895
QY	8402	AATCGTATGCTTTACGAAGCAAGTTATGAAAAATTAATGATTAACAAAAAATATA	8461
DB	3896	CTAAAGAAAAAGCTATTAATATATCTTAATTTGCAAAACAATCAATGATTTACGCAA	3955
QY	8462	TTTCAGAAATGATGCTCTTCTTTAATGATTAACAAGAAATGAAGCGAATATAGATCTT	8521
DB	3956	TTAAAGATCAAGCAGTGTCTATGTTCAAGTATTAATCTGAGATACAAATTAAGATG	4015
QY	8522	TAGCGGTAGCCGGTGTGATGCAAGAGAACAAAGCACTTTTACAGATCAACAAAGT	8581
DB	4016	TTGCGAAAGATGAAATTTAGCAACAAAGCAAGGAAACAAAGCACTTTTTCACAAATG	4075
QY	8582	TAACTTACAACTGTAAATGAGAGAAACGATATCACTTCGTGCAAAAGCTTTGGCTA	8641
DB	4076	CAGATGCACTACGAAAGAAAAAGAACAGCAAAATCAAC-----AATATATGCAAGAT	4129
QY	8642	AAAAATGAAATTTATGAAATGTAAGAAAGCTGAGAGGCTTATGTCGAGCGGAAACAG	8701
DB	4130	TAAAGCAAGTATATCAAAATATTTGAAATGACAGTCAATGATGATGTAACACTGCA	4189
QY	8702	CAGCGGTGAAATTTATACAAAGATATACAGAGCATGTTGTTGACGAAATTTGGGAA	8761
DB	4190	AAAGTAAATGCAATTTCAAGCAATTTGACCAATTTCAAGCATCAACAGATGTTAAAAAGATG	4249

```

OY 8762 TTGGAGATTAATAGAAAGATTGCAAGAGATTAATGATTGTAAGTCAACGAGACG 8821
DB 4250 CAAGGCGGAAATTTGTAAGTGAATGCAAAATTAATTAATTAATTAATTAATG 4309
OY 8822 GAACCAAGAGAGGTCTTGTGCGAAAGAAATGATTTCTGTGAAATAATCAATTTTCAGGAG 8881
DB 4310 AGACTACTAATTAAGAAAAAGTTAAGATTTGACAGCTTGAAGTGAAGTGAAG 4369
OY 8882 AAACAAAATCATCATTTGAAAGTAAAGCCAAATTTGTTGAAACCGAAGTGAATGTAAG 8941
DB 4370 GTTTAAATTAATTAATTAATGAGAACTACTAGAGTGAATGTAATCTAGTAAAGATACAG 4429
OY 8942 ATGCTTTGAATGAATCTGATGATGATCTAGCAAGAAAAAGTGGCTATGTTGGAATTTG 9001
DB 4430 CAGTCAAAAAGTTCAACCAATCTTCATGCAAACTCTTTTAAAGAGCCAGACGATTAAGAC 4489
OY 9002 GTATTGGAATTTGATGTAATTAATGATTAAGAAAAATGTAAGCCAAATCGGAA 9061
DB 4490 CTTTGAATCAAGCTGAGCTGATTAAGAAAAACAAATTAACAAACCAATTAATGCTGAC 4549
OY 9062 GACATGCTATTGTAGAACTACTGCAAAAACAAGAAATATCAAGCTTTACAGAGCAAAAG 9121
DB 4550 AACAAGAAATTAACA---TGCAAAAACAAGAAAGTTGATCTGAATTAATCAAGGCAAAA 4606
OY 9122 TAAATATTCTTGAAGAAAGAGACGCTGACGCTGCAATATGCAATGATACACATTT 9181
DB 4607 CAAATATTGACCAATCTTCACACAGATGAATGTTGATATGATGATTAAGAGAGCAAAAG 4666
OY 9182 CCAATGAGATGATTAATTAATTTGCAAGACAGATATGATCTTCTCAATTAATTAACA 9241
DB 4667 CTAAATTAATGACGTTAAACATTTAGTGAATCAAAAAGATGCTTTAGCTTAAGATTTG 4726
OY 9242 AAAATTTCAAAAATTAATTAATTTAGCTTTAGCATC-----AAGTGAATGGAATGTAAGT 9295
DB 4727 AAGCTGCAATTAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4786
OY 9296 TTCATGGGTGCTGAGCAAGAGTGCAGAGCCAAAGCAGCAGTTAGTGAAGATTC 9355
DB 4787 AATTTGCAAGAACGAAACAAAACCTGCTGAATTTGAAACAAAACGCGGAGTCAAAATGTC 4846
OY 9356 AATTAATTAATTAATTAATTTAGTGAATTTAGCAGGAAATTTAAACAGAGGAAACATCA 9415
DB 4847 ATCAAGCTACTTCTAAAGATGACATTAAGTCAATTAATTAATTAATTAATTAATTAAT 4906
OY 9416 ATGTAATATCCGGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9475
DB 4907 ACGATTACCAATTCACACAGTAAAGAAATCAGCTACACAGATTTATATGCTATG 4966
OY 9476 CGGATGCAAAAGTCAATGCTGAGCTGCTCGGCAACTGCACTATTGAAAAAATGAAG 9535
DB 4967 CAGATCAGAAAGAAATTAATTAATTTAGCTGACGCTAATCAACACAGATGAAGAAAGCAAC 5026
OY 9536 TAAATTTTAATTAATGCGATCCGAAATTTAAATTAATTAATTAATTAATTAATTAAT 9581
DB 5027 AAGCAATTAAGCAAGTTGAACAAAATGTTCAAACTCATTAAGAAAG 5072

```

RESULT 26
AAV74374
ID AAV74374 standard; DNA; 8155 BP.

AAV74374;

16-MAR-1999 (first entry)

Staphylococcus aureus contig SEQ ID #63.

Computer readable medium; vaccine; S. aureus infection; immunodetection;
cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
skin infection; surgical wound infection; scalded skin syndrome;
toxic shock syndrome; ds.

OS Staphylococcus aureus.

XX Key Location/Qualifiers

FT misc_feature 1321..1380

FT /tag= a

FT /note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

FT 3121..3180

FT /tag= b

FT /note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

FT 4921..4980

FT /tag= c

FT /note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

FT 6721..6780

FT /tag= d

FT /note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

EP786519-A2.

30-JUL-1997.

07-JAN-1997; 97EP-00100117.

05-JAN-1996; 96US-0009861P.

(HUMA-) HUMAN GENOME SCI INC.

Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA,

WPI; 1997-374922/35.

polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT scored on computer readable medium and used in the production of anti-
PT S. aureus vaccines.

Claim 1; Page 482-487; 3271pp; English.

PS This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S. aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S. aureus infection. The
CC S. aureus in a sample, S. aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S. aureus DNA sequences contained on the computer
CC readable medium

XX Sequence 8155 BP; 3300 A; 1279 C; 1452 G; 1876 T; 0 U; 248 Other;

Query Match 0.8%; Score 73.6; DB 2; Length 8155;
Best Local Similarity 40.2%; Pred. No. 0.00055;
Matches 972; Conservative 0; Mismatches 1404; Indels 42; Gaps 5;

QY 3782 TTCAAGTAAAAAAGTCTTTGGTTAAAAATATGATATTGCAAAATGCCAATAAATTGA 3841
Db 5651 TTAATAAAAAGCACTGCAAGCTAAATGAATCAGCAGCAATTTTAATATACAAATTTGCAAG 5710
QY 3842 AAGTAATGCTTTGAGTGGAGAACTCAAGTAGCAGCAGAGCAGGTTTGGAAAGATTA 3901
Db 5711 AGATTCAAGCTAGCCAGATGCAAGATGAAGAAAAACAGCAGCTGATGCTGAAGCA 5770
QY 3902 AAGAAAGTGGAGCAAGAAAAAGTTATCTATTGGGAACTTCTGCTTTCTATCAACTTAG 3961
Db 5771 ATACTGAAATAGTAAAGCAATTCAGAGCATTTTCAGAGCAACTATCAAGCAGAACTTG 5830
QY 3962 TGAACAATGAATTTCTGCAAAATCAGAAATTAATACAGTAGCAGAGAAATCTGAAGCC 4021
Db 5831 ATGAAGCTAAAGCAAAATGCAAGAGCAGCATTTAATGCCGTAAACACAAAGTTGTGAAGA 5890
QY 4022 AAAAAATGGAATGTCATGTCATGCTTATCAAGCCGACACCCCAATGACAGAGCTTTAA 4081
Db 5891 AACAGCCGCTAAAGATGAATGAATCAATTACAAAGCAACGCAACCAATGTTATCAATA 5950
QY 4082 ATTTCAAGCTGGAAGTCAAAATGGAACGTAGGGGCTACTGTGACTGTTGCCAATTGA 4141
Db 5951 ATGATCAGAACGCTACACAGAGAAAAAGAAAGCAGCTATTCAACAATTAGCAACGAG 6010
QY 4142 ACAACAAAGTAAATGCTTTATTAAGTGGGAGATATACCTAACGTTAATCAGCGAG 4201
Db 6011 TTACAGACGCGAAAAATATATTTACAGCTGCAACTGATGATTAATGTTAATCAGCGGA 6070
QY 4202 CAAAAGCTCTTTAGCAACCACTCAAGTAGCTGCTGCACTGACGAGCGGAGGACAAATTA 4261
Db 6071 AAGAGCTGGAAGAAATTCAAATTCAAAGCAGCAACAGCAACACGGTTAATCAATG 6130
QY 4262 GTTTCGAGCGGGAATTAGAAATTAATCAAGGGGCTGTTCTGCTCAATAGATTGCAATG 4321
Db 6131 CTAAAAATGATTTTATCATAGCTGTGACAACTCAAAATCAGCAATTTGTAATACACTG 6190
QY 4322 ACGTGAAGCTAGCGTTATTAATCTTCATCGAAGAGCTAATGAAAATCAATGTCATG 4381
Db 6191 GTGCTACAACTGAAGAAAAATGCAAGAAAAGTTTGGTTTAAAGCTAAAGAAAAAG 6250
QY 4382 CCAAGATGTCAAAGAAATTTGATCTAGCAAAAAGAAATATCAGGCTTTACTAAATGAA 4441
Db 6251 CGTATCAAGATATCTTAATATGCAACAACAACTAATGATTACGAAAATTAAGATCAAG 6310
QY 4442 AAGATAAAAAATTTAGAAAGATCGTGTATTATACGCTGGAATGTTATTAAGA 4501
Db 6311 CAGTTGCTGATATTCAAGGTATTACTGAGATACAACTTAAGATGTG-----CGA 6364
QY 4502 AAGAACAACTGAAAAAGCAAGAAAAAGAAAGAGCGGTCATTTGTAATGCTGCTTAT 4561
Db 6365 AAGATGAATTTAGCAACAAAAGCAACGAAAAAGCGTTATTGCAAAACTGCAGATG 6424
QY 4562 CGGTTGCTGGAACGATTAATCCGCTGAGAGATGCTATTGCACTCAATCTGTTAAA 4621
Db 6425 CGACTACTGAGAAAAAGAAACAAAGCAAAATCAACAAGTAGACGCAATTTAACACAAGGTA 6484
QY 4622 ATAAATTTAAAGCAGAAATTTAGTGAAGCAATTAAGAAAGCCGAGAGAGATTAATTCATG 4681
Db 6485 ATCAAAATTTGAAATGCAACGATCAATCGATGATGAATCACTGCAAAAGTAATTCGA 6544
QY 4682 CGAAACATGTAATTTGAGAGCAAAATCATCTACTGTTGTGTGTAATCGGCTTTGAGC 4741
Db 6545 TTCAAGCAATTTGACCAATTTCAAGCATCAACAGATGTTAAAAAGAAATGCAAGAGC-GGAA 6603
QY 4742 TTGCTATGACAAAGATGCTTTTTCAGGAATGGATCTGAGCATGGAAGACTTATCA 4801
Db 6604 TTGCTAACTGAAATGCAAAAATAAATTAATCTGAATTAATTAATAGAGACT----- 6657
QY 4802 ATGACACGATTTGCAAAAGGTGATTAAGAGAAATTTCTGCTGATTTCTTAAATGTGAACG 4861
Db 6658 -----ACTAATGAAAAAAAGGTAAAGATTTGACAGCTTAGACACACATATGAAGAG 6712
QY 4862 CAAATATTCATCTTGGGGTGAATGTGCGGAACAATGCCGGTTCTCTTTCTACGG 4921

Db 6713 GTTTAAATNN 6772
QY 4922 CGGTAGAGCGTCTTTTGCAGATATACTCTTCAATTAATAAACCTGCTTGTGATTACAG 4981
Db 6773 NNNNNNNNAAAGTTCAACAACTTCAATGCAATCTCTTTAAAGAAACAGAGTAAAAAG 6832
QY 4982 GAACGAAGTAAATCTTTTATGAGAAAGAAATCAAAAGTCAATGTACAACTTTGAATG 5041
Db 6833 AATTAGATCAAGCTGACAGCTGATTAAGAAAAACAAATAGAACAAACCAATATGATCAG 6892
QY 5042 ATTCTATTTTACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAGAGGCTGGAATCG 5101
Db 6893 AACCAAGAAATTAAGATGCAAAAACAAGAGTTGATCTGAATTAATCAACGAAAAACA 6952
QY 5102 GAGGAATGATCTGTCAATGCTGTTCTGATGAAACGGAAGCTTATGATGATTCG 5161
Db 6953 ATGTGATCAATTCATCAACAAATGAATATGTTGATTAATGACGTTAAAGAGAAAAAGCTA 7012
QY 5162 AGTTGAAGAGTAAAGTTCTTTCAATGTAGATGCAAAAAGATCAAAAAACAATTAATACAA 5221
Db 7013 AAATTAATGCACTTTAAACATTTATGATGATCAAAAAAGATGCTTTAGCTAAATTTGAAG 7072
QY 5222 TTGCCGGAATGCAATGAGAGAAAAAGCGCTGAGATTGAGCAACAGTTGCTCATACAA 5281
Db 7073 ATGCAATTAATGCTAAAGTAAACGAAGCGGATTAACCTTAACGATCGACTTCAAGTAAA 7132
QY 5282 ATATTGAAAACAATCACTTATAGCTATTGTAATAAACAGTAAATTAACAACGCGAATG 5341
Db 7133 TTGCTGAAGCGAAACAAAACTTGTGATTTAAACAACTCGGATTCAAATGTTAATTC 7192
QY 5342 ATCAAGATAGAAAAATATCAATGTGACTGCAAAAAGATTAATGATGCAATTAATAG 5401
Db 7193 AAGCTACTCTAAAGATGACATTTGAAGTTCAATTTCAATGATGATTAATTAACG 7252
QY 5402 CAGTCGAGATTGAGAGCAAAAGAGAGCTGTGCAAGAGCTTTCGCAAGTACTACT 5461
Db 7253 ATTAACAATTTCCAAACAGGTAAAGAAAGATCAGCTAACACAGATTTATATGCTTATGAG 7312
QY 5462 TGAATTAAGACGTTTCTTCTCATGTTGATCAAACTGATTAATGACAAAGATTTAAGAGAG 5521
Db 7313 ATCAGAAAGAAAAATTAATTTTACAGCTGACCTAATGCAACAAAGATGAAAAAGCAACAG 7372
QY 5522 AAAATTAATGGAATTAAGAAAAAGCAATGTTAATGTTCTAGCTGAAAAATACAGTCAAG 5581
Db 7373 CAATTAAGCAAGTTGACCAAAATG---TTCAAACCTGCTTGAAGCACTTAATTAATGTTG 7429
QY 5582 TGTTCACAAATGCGACAGTGTCTTTCGAGAGCAAGTGAACAAGTGCAGTAGAGCTGAG 5641
Db 7430 TGAATTAATGAGAGCTGATGATGATTAACAAGGTAAGAGCAAGCAATGATGCTATTC 7489
QY 5642 TAGCAGTTAATTAATTAACAAATTAATCTTGCACATATTAATAATTAATGCTCAAAATG 5701
Db 7490 AAGTAGATGCTAAGCTTAATTAAGGCAACCAAGCTAATGAAGTTAAAGCAGAAAGTAA 7549
QY 5702 TAGGAATGCTTGTGTAAGAG-----GAAATCTGATTCATCTA 5740
Db 7550 CGAAAGAAATCTAATGATCAAGAGTACCAAGTTAATCTGTAAGAAAAAATCAAGAGATTAG 7609
QY 5741 TTAATAACAATTTGAAATTTGAGAGCTGAGAGTGTGAGAGCTGAGAGTCAAGTTTCA 5800
Db 7610 CAATGATTTAAACAAATTAACAGATCAAGCTAAACAAGATTAATCTGATGCAACCAACTG 7669
QY 5801 TAGCAGTGAATTAAGATTGTAATAATACATAGACAAATTAATCATGCAAAAAATCACTG 5860
Db 7670 CTGAAGTTGAAAAAGCAAAAGCTCAAGGACTTGAAGCACTTGAATTAATCATTAATGCACT 7729
QY 5861 CGAAGGAAATGTGAGATTATTAACAAGTGTGATGCGGTAAATGCTAATTAATGCAAGAA 5920
Db 7730 CAACAGAAAAACAAGAAAGCTATGAGAAATTAAGAACTGCACTGACCAAGATTAAGAGAG 7789
QY 5921 CAGTGTCTGAGGTGCGCGGTGACAGCAATAGGAGCTCAACAGGTGTAATGAATTTACAG 5980

Db 7790 GGTGAATGTCACGCTGATGCTACAGCAAGAAAAGAGCGTTTACGAATGCTTAG 7849
Qy 5981 GATTCACAAAGCATATGTAAAGATTCTACAGTATGCTTAAGAGAAACAGATGATT 6040
Db 7850 AAGACATTTTATCAAAAGCACTGAAAGATTTTCATCAAACTCAAAATGCAAAATCG 7909
Qy 6041 AATTACTACTCAAGCGCAAGTAAAGTGTGATTAAGTAAATTCAAAATCTTAATA 6100
Db 7910 CTACTGTCAAAAATAGTGGCTTGACCACTTAAGCACACAGTATTAATCTGAAGTTA 7969
Qy 6101 TTACGAGAGACTTATACCAAAAAGAAAATAAGTAAATTAAGATTTGTTACCAATA 6160
Db 7970 AAAAAAATCTTTCGAAAGCAATCAGAAAGTGTTAACAGCAAAATAGAAATTAATAA 8029
Qy 6161 GTTCAGCTACTCAATCTT 6178
Db 8030 ATGCAGATGCAATGATGAT 8047

RESULT 27
ACAI9708
ID ACA19708 standard; DNA; 7446 BP.
ACAI9708;
19-JUN-2003 (first entry)
Prokaryotic essential gene #1365.
Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
OS Staphylococcus aureus.
PN MO20027183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342823P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELI-7) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haseibeck R, Ohlsen KJ, Zykind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX MPI; 2003-029926/02.
XX P-PSDB; ABUS838.
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 7578; 1766p; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the isolated
CC nucleic acid; (2) a host cell containing the vector; (3) an antisense
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a gene that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 7446 BP; 3198 A; 1311 C; 1311 G; 1625 T; 0 U; 0 Other;

Query Match 0.7%; Score 72.4; DB 8; Length 7446;

Best Local Similarity 41.8%; Pred. No. 0.0092;

Matches 804; Conservative 0; Mismatches 1091; Indels 27; Gaps 5;

Qy 3782 TTCAAGTAAACCAAGTCTTTGTTAAATAGTATTTGGAATGCAATTAATTAA 3841
Db 3608 TTAATTAATAAACCACTGACCTAATTAATGAATCAGCAATTTTAATCAAAATGCAAG 3667
Qy 3842 AAGTAAATGCTTTGAGTGAAGAACTCAAGTACAGCAGACAGAGTTTGAAGCAATTA 3901
Db 3668 AGATTCAGCTACACCAATGCAACAGATGAAGAAAACAGCAGCTGATGCAAGAA 3727
Qy 3902 AAGAAAGTGAGAGCAAGAAAAGTATCTATTGGGAACTTTCCTTCAACTTNG 3961
Db 3728 ATACTGAAAATGTTAAGCAAAATCAAGCATTTGACAGCACTAATGCAAGAGTTG 3787
Qy 3962 TCAACAAATGATTTCTGCAAAATCAAGAAATTAATACAGTACAGAGAAATGCAAGCC 4021
Db 3788 ATGAAGCGAAAGCTAATGCAAGAGCAATTAATGCGTTAACCACAAAGTTGTTAAGA 3847
Qy 4022 AAAAAATGATTTGATGTCATGCTTTATCAAGCGGACACCCAGTGAACAGAGCTTTAA 4081
Db 3848 AACCAAGCGCTAAAGTATGATTCATTAACAGCAACGCAACAAATGTTATCAATA 3907
Qy 4082 ATTACAAAGCTGAAAGTCAAAATGGAATCTGAGGAGCTACTGCTGCAATTTAA 4141
Db 3908 ATGATCAGAAACGCTCAAAATGAAGAAAAGAGCAGCTATTCAACATTTGCAACAGCAG 3967
Qy 4142 ACACAAAGTAAATGCTTTATTAAGTGGGAGATTAATCTTAACGTTAATGACGGAGAG 4201
Db 3968 TTACAGACCGCAAAATTAATTAATTAAGCTGCAACTGATATATGTTAGATACAGCA 4027
Qy 4202 CAAGAAGCTTTTACCAACCACTCAAGTGAAGTCTGCAAGTGAACGAGAGGACAAATTA 4261
Db 4028 AAGAGCGCTGAAGAAATTAATTCAAAGCAACACACGCAACAGAG-----TTA 4078
Qy 4262 GTTCTGAGCGGAGATTAGAAAATTAACAGGAGGCTGTTCTGCAATAG---ATTGACA 4318
Db 4079 AATCGAATCGCAAAAATGAATGAATGATCAAGCTGTGCAACCTCAAAATCAAGCAATGATA 4138
Qy 4319 ATGACGTGAAGACTAGCTGATTAATCTTCATGCAAGAGACTATGAATCAATGCTCA 4378
Db 4139 ATACAACTGGCCCTCACTCAAGAGAAAATGACGCAAAAGATTTAGTTTAAAGCTTA 4198
Qy 4379 TTGCCAAGATGTCAAAGGAAGTTCTGATCTAGCAAAAGATATCGCTTTACTTAATG 4438
Db 4199 AAGAAAGACGATCAAGATATCTTAATGCAACAAACATATGATTTAGCAAAATTA 4258
Qy 4439 GAAGATTAATAAATTTTAAGAGATGCTGTTAATTAACGACTGGAATGTTATTATA 4498
Db 4259 AAGATCAACAGATGCTGATTAATTAAGAGTATTACTGCAATATCAACAAATTAAGATGTTG 4318

4499 CGAAGAACACTAGAAAAAGCAAAAGAAAAAGAGCGGCTCATTTGTAATGCTGCTT 4558
 4319 CAAAAGATGAATATAGCAACAAAAGCAAAAGCAAAAGCGCTTATTGCAACAACCTGAG 4378
 4559 TATCGGTGCTGGAAACGATTAATCCGCTGGAGAGTACTATTTCAGTCATTAATCTGTTA 4618
 4379 ATGCACTACTAAGAAAAAGAAAGCAAGCAATCAACAGTAGACACCAATTTAACACAG 4438
 4619 AAAATAAATTTTAAAGCAGATTGAGTGAAGCAATTAAGAACCCGAGAGATTAATTC 4678
 4439 GTTAATCAAAATATTGAAAAATGCAAGCTCAATGATGATTAACACTGCAAAAAGTAAATG 4498
 4679 ATGCAAAACATGTAATGAGGCAAAATCACTACTGTTGTTGTAATGCGGCTTCTG 4738
 4499 CAATTCAAGCAATTGACCAATTTCAAGCATCAACAGATGTTAAACGAATGCAAGAGC-A 4557
 4739 GACTTGCTATCAGCAAAAGTCTTTTCAAGAAATGGATCTGGAGCATGGCAAGCTTAT 4798
 4558 GAATTGCTAATGCAAAATGCAAAATTAATTAATGTAATGTAATGTAATGTAATGTAATG 4614
 4739 CAATGACACAGATTGCAAAAGTGGATTAAGAAAGAAATTTCTGCTGATTTCTTAATGTA 4858
 4615 -----ACTAATGAAAGAAAAAGTACGATTAATGACACGATTAGACAGCATATGAA 4666
 4859 ACGCAATTAATTCATTTCTTGGGAGTGTGCGGAAACCAATGCCGTTCTCTTCTTA 4918
 4667 AAGTTTAAATTAATTAATTAATGCAAGCACTACTACAGTATGTAATTAATTAATGAA 4726
 4919 CGGCGTACAGAGCTGTTTGGGAATTAATCTTCAATTAATTAATTAATTAATTAATTAAT 4978
 4727 CAGCAGTACAAAAGTTCAACCACTTCAATGCAAAATCTGTTAAGAAACAGAGGTAA 4786
 4979 CAGAACAGAAAGTAAATCTTTTAAAGTGAAGAAATTAACAAATGCAATGACAACTTGA 5038
 4787 AAGAAATTAATCAAGCTGACGATGATTAAGAAACCAAAATTAACAAACCAAAATGAT 4846
 5039 ATGATTTCTATTAACAAAGTTTCTGCTGAGCGCTGCAAGTAAAGCAAGCTGAA 5098
 4847 CACAACAAGAAATTAATGATGCAAAACAAAGAGTTGATCTGAATTAATTAATCAAGGAAA 4906
 5099 TCGAGAGAAATGATTTCTGCAATCGGTTCTGATGAAACGAAAGCTTAAATGATGAT 5158
 4907 CAATGTCATCAATCATCAACAAAGATGATGATGATGATGATGATGATGATGATGATGAT 4966
 5159 CTGAGTTTGAAGAGTAAATCTTTCAATGATGATGATGATGATGATGATGATGATGATGAT 5218
 4967 CTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5026
 5219 CAATGCGGAAATGCAAAATGAGAAAGCGGCTGAGTGGAGCAACAGTTGCTCAT 5278
 5027 AAGATGATATATATGCTAAAGTAAAGCAAGCGGATTAATCTTAACGATGATGATGATGAT 5086
 5279 CAATTAATTTGGAACCAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 5338
 5087 AAATGTCGAAAGCAAAACAAATCTGCAATTAATTAATTAATTAATTAATTAATTAAT 5146
 5339 ATGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5398
 5147 ATCAAGCTACTTCTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5206
 5399 TAGCAGTGGAGTGGAGAGCAAAAGAGCTCTGTGTGCAAGAGCTTGTGCAAGTACTA 5458
 5207 ACGATTAACCAATTCACAGGTAAGAAAGAAATCAAGCTACCAAGATTAATTAATGCTTAAT 5266
 5459 CTTGAATTAAGACAGTTTCTTCTCATGTTGATCAACCTGATTAATTAATTAATTAATTAAT 5518
 5267 CAGATCGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5326
 5519 AAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5578
 5327 AAGCAATTAAGCAAGTTGACCAAAATGTTCA---AACTGATTAATGAAAAATTAATTAATG 5383
 5579 AAGTGTCACAAAATGCAAGTGTCTTTCCGAGCAAGTGGACAGCTGCAAGTGAAGGCTG 5638

Db 5384 GTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5443
 QY 5639 GAGTAGCAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5698
 Db 5444 TTCAAGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5503
 QY 5699 AT 5700
 Db 5504 AT 5505

RESULT 28
 ACF05848
 ID ACF05848 standard; DNA; 7446 BP.
 XX
 AC ACF05848;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Methicillin resistant Staphylococcus aureus ORF SA1964-fmb (mrp).
 XX
 KM Vancomycin; teicoplanin; MRSA; diagnosis; SA1964; mrps; fmb;
 KM antibiotic resistance; gene; de.
 OS Staphylococcus aureus.
 FH Key location/Qualifiers
 FT CDS 1..7446
 FT /*tag= a
 PN W02003062466-A2.
 XX
 PD 31-JUL-2003.
 XX
 PF 16-JAN-2003; 2003WO-GB000129.
 XX
 PR 16-JAN-2002; 2002GB-00000865.
 XX
 PA (UNBR-) UNIV BRISTOL.
 XX
 PI Welsh TR, Avison MB, Howe RA;
 XX
 DR WPI; 2003-636743/60.
 DR P-PSDB; ABR62804.
 PT Detecting a strain of Staphylococcus aureus having reduced susceptibility
 PT to vancomycin and/or teicoplanin, comprises detecting the presence of
 PT mutation(s) leading to loss of function of at least one gene.
 XX
 PS Claim 7; Fig 3; 78pp; English.
 XX
 CC The present sequence is that of open reading frame ORF SA1964-fmb (mrp)
 CC 2218122-2225567 of methicillin resistant Staphylococcus aureus (MRSA)
 CC strain N315. The ORF was identified as an mrp/fmb gene. The invention
 CC provides a method of detecting a strain of Staphylococcus aureus which
 CC has reduced susceptibility to vancomycin and/or teicoplanin by detecting
 CC the presence of one or more mutations in at least 2 open reading frames,
 CC including the present sequence. Polypeptides, polynucleotides,
 CC antibodies, probes, primers, kits and methods are provided for this
 CC detection. The methods are rapid, reproducible and sensitive, and allow
 CC for detection of strains that are fully or heterogeneously resistant
 XX
 SQ Sequence 7446 BP; 3198 A; 1312 C; 1311 G; 1625 T; 0 U; 0 Other;
 Query Match 0.7%; Score 72.4; DB 10; Length 7446;
 Best Local Similarity 41.8%; Pred. No. 0.00092;
 Matches 804; Conservative 0; Mismatches 1091; Indels 27; Gaps 5;
 QY 3782 TTCAAGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3841
 Db 3608 TTAAATTAAGCAAGCTGCAAGTAAATGAAATCAAGCAATTTTAATTAATTAATGCAAG 3667

QY 3842 AAGTAATGCTTTGAGTGAGGAACCTCAAGTACGACGACGACGCTTTGGAGCAGTTA 3901
 Db 3668 AGATTCAAGCTACACACCAATGCAACGATGAGAAAAACAAGCAGCTAATGACGAGGAA 3727
 QY 3902 AAGAAAGTGAGGACCAAGAAAAAGTTATCTATTGGGAACTTCTGCTTCTATCAACTAG 3961
 Db 3728 ATACTGAAAAATGTTAAAGCAAAATCAAGCATTTCAGCAGCACTACTAAGCACAAGTTG 3787
 QY 3962 TGAACAAATGAGTTTCTGCAAAATGAGAAATATATACAGTACAGAGAAATCTGAAGCC 4021
 Db 3788 ATGAAGCAAAAGCTAATGACAGAGCAATTAATGCGGTAAACCAAAAGTTGTTAAGA 3847
 QY 4022 AAAAAATGATGTTGATGCTACTGCTTATCAAGCGGACACCCAAATGACAGAGCTTAA 4081
 Db 3848 AACCAAGCGCTAAAGATGATGATTCATATTACAGCAACGCAACAAATGTTATCAATA 3907
 QY 4082 ATTTTCAAGCTGGAAGTCAAACTGTAGGGGCTACTGTGACTGTGCAATTAA 4141
 Db 3908 ATGATCAGAACGCTTCAAAATGAGAAAAAGACAGCTATTCAACAATTAGCAACGAG 3967
 QY 4142 ACAACAAATGAATGCTTCTATTAAGTGTGGAGATATATCACTTAATCGAGCGGAGC 4201
 Db 3968 TTACAGACGCGGAAAAATATATTAAGCTGCAACTGATGATTAAGTGTAGATACAGCGA 4027
 QY 4202 CAAAAGCTTTTAAACCACTCAAGTACTGTGCACTGACGACGCGGAGGAGCAATTAA 4261
 Db 4028 AAGACGCTGGAAGAATTCATTCAAAGCACAAACAGCAACAGAG-----TTA 4078
 QY 4262 GTTTCGAGCGGGATTAGAAATTAATCAAGGGGCTGTTCTGCAATAG---ATTGACA 4318
 Db 4079 AATGCAATGCAAAAAATGAAAGTTGATCAAGCTGTGCAACTCAAAATCAAGCAATTGATA 4138
 QY 4319 ATGAGCTGGAAGCTAGCTTCTATTAAGTGTGGAGATATATCACTTAATGAATCAATGTA 4378
 Db 4139 ATACAACTGCGCTTCACTGAAGAAAGAAAAATGACCAAAAGATTAGTTTAAAAAGCTA 4198
 QY 4379 TTGCCAAAGATGTCAAAGAGTTCTGATCTACCAAAAGAAATTCAGGCTTTACTAATAG 4438
 Db 4199 AAGAAAAACGCTATCAAGATATCTTAAATGCACAAACTAATATGTTAACCAATTA 4258
 QY 4439 GAAAAGATAAAAATTTTGAAGATCGTGTATTAATGCACTGGAATGGTTATTATA 4498
 Db 4259 AAGATCAAGCAGTGTGATTAATTCAGATTAATCTCAATACACAAATTAAGATGTTG 4318
 QY 4499 CGAAGGACAACCTAGAAAAAGCAAAAGAAAAAGAGAGCGGTCAATGTAAATGCTGCTT 4558
 Db 4319 CGAAAGATGAATTAAGCAAAAAAGCAAAAGCAAAAGCGCTTATTGCAAAACGCGAG 4378
 QY 4559 TATCGGTGCTGGAAGGATTAATTCGCTGAGAGAGTACGATTTGCAATCAATACGTTA 4618
 Db 4379 ATGCGCACTAGTGAAGAAAAAGCAAGCAAAATCAAGTACGACCAATTTAACCAAG 4438
 QY 4619 AAAATTAATTTAAAGCAATTAAGTGAAGCAATTAAGAAAGCGGAGAGGATTAATTC 4678
 Db 4439 GTAAATGAAATATTTGAAGATGACAGCTCAATGATATGTAACACTGCAAAAGATTAAG 4498
 QY 4679 ATGCAAAACATGTAATGTTGAGGCAAAATCATCTACTGTTGTGTAATGCGCTTCTG 4738
 Db 4499 CAATTCAAGCAATTAAGCAATTCAGCACTCAAGATGTTAAAGCAAGCAAGC-A 4557
 QY 4739 GACTTCTATCAGCAAAAGATGCTTTTTCAGGAAATGGAGTGTGAGCATGGCAAGCTTAT 4798
 Db 4558 GAATTCCTAATGAAATGCAAAATTAATACTGAATATCTTAATATATATGAGACT--- 4614
 QY 4799 CAAATGACGATTTGCAAAAGTGTATTAAGAGAAATTTGCGTGAATTCCTTAATGTTGA 4858
 Db 4615 -----ACTTAATGAGAAAAAGTAAAGATTAATTTGAGCAAGTTAAGCAGCTATGAGAG 4666
 QY 4859 ACCCAATTAATTCATTTCTGGGGTGAATGTTGGGGAACCATTCGCGGTTCTCTTTCTA 4918
 Db 4667 AAGTTTAATTAATTAATTAATGACAGCACTACTACAGTGTATGTAATGCTTAAGAGATA 4726
 QY 4919 CGGCGGTAGAGCTGCTTTTTCGCAATTAATCTTCTTAATTAATAAACCCTGCTTGTATTA 4978

Db 4727 CAGCAGTACAAAAAGTTCAACACTTCAATGCAAAATCTGTTAAGAAACACAGAGTAAAA 4786
 QY 4979 CAGAAAGCAAGCTTAATCTTTTGTGAAAGAAATACAAAGTCAATTAACAGTTTGA 5038
 Db 4787 AAGATTAATGATCAAGCTGACAGCTGATTAAGAAAAACAATATGAAACAAACCAATATGAT 4846
 QY 5039 ATGATTCATTAATTAACAAGCTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGGA 5098
 Db 4847 CACAACAAAGAAATTAATGATCAAAAAACAAGATTTGATCTGAATTAATCAAGGAGAAA 4906
 QY 5099 TCGGAGAAATGATATCTGTCATGCTGTTCTGATGAAACGGAAGCTTTAGTTAGATT 5158
 Db 4907 CAATGTCATCAATCAACAAATGAAATGTTGATTAATGACAGTTAAAGAGAAAAAG 4966
 QY 5159 CTGAGTTGAAGAGTAAGTTCTTCAATGTAATGATCAAAAGATCAAAAAACAATTAATA 5218
 Db 4967 CTAAATTAATGAGCTTAATTAACATTTAGTGAATCAAAAAAGATCTTTAGCTAAATTTG 5026
 QY 5219 CAATGCGGAAATGCAAAATGAGAGAAAGGCGCTGAGGTTGAGCAACAGTTGCTCATA 5278
 Db 5027 AAGATGCAATTAATGCTAAAGTAAAGAGGAGTAACTTAACGCAATGCACTTCAAGTG 5086
 QY 5279 CAATATTTGAAAAACAATCAGTTATAGCTATTTGTAATAAACAAGTAAAAATTACAGCGGCA 5338
 Db 5087 AATTTGCTGAAGCGAAACAAAACTTGCTGAATTAACAAACCTGCGATCAAAATGTTA 5146
 QY 5339 ATGATCAAGATTAAGAAAAATATCAATGTAATGCTGCAAAAGATTAATCTATGACCAATCTA 5398
 Db 5147 ATCAAGTACTTCTTAAGATGACATTTGAAGTTCAATTCATTAATGATTAATTAATTA 5206
 QY 5399 TAGCAGTGGAGTTGAGAGAGCAAAAGAGCCTCTGTGCAAGAGCTTGTGAAGTACTA 5458
 Db 5207 ACGATTAACAATTTCCACAGGTAAAGAAATCAAGTACACAGATTTATATGCTTAAG 5266
 QY 5459 CCTGAATTAAGACAGTTTCTTCTCATGTTGATCAATGATTAATTTGCAAAAGATTTAGAG 5518
 Db 5267 CAGATCAGAGAAAAATATATTTTCAAGTGAACATTAATGCAACAGATGAAAAAGCAAC 5326
 QY 5519 AAGAAATTAATGAATAAGAAAAAGCAAAATGTTAATGTTTGTAGTCAAAATACAGATTC 5578
 Db 5327 AAGCAATTAAGCAAGTGAACCAAAATGTTCA---AAGTCAATTAAGAAAAATTAATTAATG 5383
 QY 5579 AAGTGTCAAAATGCGACAGTCTTCCGAGCAAGTGAAGAGTGAAGAGCTG 5638
 Db 5384 GTGTGATTAATGATGACGTTGATGATGATTAATCAACAGTAAAGCAGCAATTTGATACTA 5443
 QY 5639 GAGTGAAGTAAATTAATTAACAAAAATTAATTTGACATATTAATAATTAATGATCTCAAA 5698
 Db 5444 TTCAAGTGAATCTACTGTTAACTTAAGCAACCAAGCTATTGAAGCTTAAGCAGAAAG 5503
 QY 5699 AT 5700
 Db 5504 AT 5505

RESULT 29
 ID ADU74046 standard; DNA; 7446 BP.
 ADU74046:
 AC ADU74046:
 XX 10-FEB-2005 (first entry)
 XX
 XX
 DB Staphylococcus aureus Msp homologue coding sequence.
 XX ANTIIDRIOTIC-RESISTANCE; VRSA infection; Msp homologue; gene; db.
 XX Staphylococcus aureus.
 OS
 XX
 XX Key Location/Qualifiers
 FH 1..7446
 FT CDS
 FT /*tag= a

/product= "Mrp homologue"

MO200409444-A2.
18-NOV-2004.
07-MAY-2004; 2004MO-GB001985.
09-MAY-2003; 2003GB-00010794.
(UVR-) UNIV BRISTOL.
Walsh TR, Avison MB, Howe RA, Wootton M;
WPI; 2004-821685/81.
Detecting a strain of *Staphylococcus aureus* with reduced susceptibility to vancomycin or telicoplanin by relating the detected expression level of *fmbB*/*mrp* or homolog within the strain to susceptibility to vancomycin or telicoplanin.
Claim 2; SEQ ID NO 17; 102bp; English.
The present sequence is the coding sequence for an Mrp-like protein of a vancomycin-sensitive *Staphylococcus aureus* strain, denoted SA1964-fmbB (mrp) 2218122-2225567. The encoded Mrp homologue protein may have a similar function to *S. aureus* Mrp/fmbB proteins. Expression of the *mrp* gene is down-regulated in *S. aureus* strains with a decreased susceptibility to vancomycin. The invention provides a method for detecting a strain of *S. aureus* which has reduced susceptibility to vancomycin and/or telicoplanin. The method involves detecting the level of expression of an *fmbB*/*mrp* gene or an *mrp*-homologue gene (including the present sequence) or a variant within the strain and relating that to the susceptibility or resistance of the strain to vancomycin and/or telicoplanin. Detection of expression is performed using an immunological method to detect the gene product or a molecular method such as RT-PCR to detect RNA constructs. The method is rapid, reproducible and sensitive and allows, for example, discrimination of vancomycin intermediate *S. aureus* (VISA) and heterogeneous vancomycin intermediate *S. aureus* (hVISA) from vancomycin/telicoplanin sensitive *S. aureus* (VSSA).
Sequence 7446 BP; 3198 A; 1312 C; 1311 G; 1625 T; 0 U; 0 Other;
Query Match 0.7%; Score 72.4; DB 13; Length 7446;
Best Local Similarity 41.8%; Pred. No. 0.0092;
Matches 804; Conservative 0; Mismatches 1091; Indels 27; Gaps 5;
3782 TTCAAAGTAAACAAGTCTTGTAAATAGTATTCGAATGCCAATTAATTGA 3841
3608 TTAATTAATAAACCACTGCACGTATGAATCAAGCAATTTAATAACAATTGCAAG 3667
3842 AAGTAATGCTTGAAGTGAAGAACTCAAGTAGCAGCAGAGACAGTTTGAAGCAAGTTA 3901
3668 AATTTCAGCTACACCAAGATGCAAGATGAAGAAAAACAGACAGTATGAGAGCA 3727
3902 AAGAAGTGAAGACAGAAAAAGTTATCTATTGGAACTTCTGCTTCTATCAACTTAG 3961
3728 ATACTGAATAATGTAAAGCAATCAAGCATTTGAGCAGCACTACTAAGCAGACAGTTG 3787
3962 TGAACATGAAGTTCTGCAAAATCAGAAATTAATACAGTAGCAGAGAAATGTGAAGCC 4021
3788 ATGAAGCCGAAAGCTAATGCAAGAGCAGCAATTAATGCGGTAAACACCAAAAGTTGTTAAGA 3847
4022 AAAAAATGATGTATGATGCTACGCTTATCAAGCCGACACCAAGTGCACAGAGCTTTAA 4081
3848 AACCAAGCGGCTAAAGATGATGATTCATTCAAGCAACGCAACCAATGTTATCAATA 3907
4082 ATTTCAGAGCTGAAAGCTCAAAATGGAATGTAGAGGCTACTGTGACTGTGCCAATTTAA 4141
3908 ATGATCAAGACCTCAAAATGGAAGAAAAAGCAGCTATTCAACAATTTAGCAACGCGAG 3967
4142 ACAACAAAGTAAATGCTTCTATTAGTGTGAGATATTAACGTTAATCAGAGCGAGC 4201

3968 TTAACAGCGCGAAAAATTAATTAACAGCTGCACATGATGATATGTTGTATGATACAGCGA 4027
4202 CAAGAAGCTTTTATGCAACCACTCAAGTACTGCTGACGTGACGACGAGGAGGACAAATTA 4261
4028 AAGACGCTGGAAAGAAATTCATTTCAAGACACAAACGACACAGCAG-----TTA 4078
4262 GTTCTGAGCGGATTTAGAAATTTATCAAGGGGCTGTTTCTGTCAATTAAG---ATTGACA 4318
4079 AATCGAATGCGAAAAATGAAGATTGATCAAGCTGTGACAACTCAAAATCAACCAATTTGATA 4138
4319 ATGACGTGAAGCTTAGCTTATTAATCTTCATCGAAGAGCTTAATGAATCAATGTCA 4378
4339 ATCAACTGGGCTTACCACTGAAGAGAAAAATGACAGCAAAAGATTTATTTAAAGCTA 4198
4379 TTGCAAAAGATGTCAAAGAAAGTTCTGATCTAGCAAAAAGAAATATCAGGCTTTACTTAATG 4438
4199 AAGAAAAAGGCTATCAAGATATCTTAAATGCAAAACAACTAATGATGTTACGCAATTA 4258
4439 GAAAAATAAAAATTTAGAAAGATCGGTATTAATACACTGGAATGATTAATTA 4438
4259 AAGATCAAGAGTGTGCTGATATTCAGAGTATTACTGACAGATACAACTTAAGATGTTG 4318
4499 CGAAGAACCACTGAAAAAGCAAAAGAAAAAGAGCGGCTCATTTGTAATGCTGCTT 4558
4319 CGAAAGATGAATTTAGCAACAAAGCAAAAGCAAAAGCGCTTATTTGACAAACTGCGAG 4378
4559 TATCGGTGCTGGAACGGATTAATCCGCTGAGAGATGATTAATGCAAGTCAATCTGTTA 4618
4379 ATGCACTACTGAAGAAAAAGAACAAAGCAATCAACAGTGAAGCAGCAATTAACACAG 4438
4619 AAAAAATTTTAAAGCAAAATGAGTGAAGCAATTAAGAAACCGAGAGATTAATTC 4678
4439 GTATCAAAATATTTGAAATGCAACAGTCAATCGATGATGTAACATGCAAAAGATATG 4498
4679 ATGCAAAATCTGAATATGAGAGCAAAATCTACTGTTGTTGTGAATGCGGCTTCTG 4738
4499 CAATTCAGCAATGACCCCAATTCAGAGTCAACAGATGTTAAACGAATGCAAGAG-A 4557
4739 GACTTGCTATCAGCAAGAGATCTTTTTCAGGAATGGATCTGAGCATGGCAAGACTTAT 4798
4558 GAATTTGCTAATGAAATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4614
4799 CAAATGACAGATTCGAAAGGTGATTAAGGAAGAAATTTCTGCTGATTCCTTAATGTGA 4858
4615 -----ACTAATGAAGAAAAAGTAAAGTATGGAACAGTTAGACAGACATATGAAG 4666
4859 ACGCAATTAATTCATCTTGGGGTGAATGTTGGGGGACCAATGGCGGTTCTTTCTTA 4918
4667 AAGGTTTAATTAATTAATTAATGACCAACTACTACAGGTGATGTAATTAATCTGCTAAAGATA 4726
4919 CGGCGTGAAGAGCTGTTTTCGAATTAATCTTCAATTAATTAATTAATTAATTAATTAATTA 4978
4727 CAGCAGTCAAAAGTTTCAACACTTCATGCAATCTGTTTAAGAAACAGAGGTTAA 4786
4979 CAGGACGAAAGTAAATCTTTTATGAGAAAGAAATACAAAGTCAATGTATCAAGCTTTGA 5038
4787 AAGAAATTAATCAAGCTGACACTGATGAAGAAACAAATGGAACAAACCAATATGAT 4846
5039 ATGATCTCATTTACAAAGCTTTCTGCTGAGAGCGCTGCAAGTTATTAAGACAGCTGGA 5098
4847 CACAAACAAAGAAATTAATGATGCAAAACAAAGAAATGTAATCGAAATTAATCAACGAAAA 4906
5099 TCGAGAGAAATGTAATCTGTCAATGCTGATGTAAGAAACGAAAGCTTTAGTGTGATT 5158
4907 CAAATGCTGATCAATCAACAAATGAAATGTTGATTAATGCTTTAAAGAGAAAG 4966
5159 CTGAGTTTGAAGAGTAAAGTTCTTTCATGTAGATGCAAAAGATCAAAAAACAAATTAATA 5218
4967 CTTAAATTTAATGCAAGTTTAAACATTTAGTGAATGCAAAAAAGATGCTTTACTAAATTTG 5026
5219 CAATTCGCGAAATGCAAAATGAGAAAGCGCTGAGATTGAGCAACAGTGTCTACTA 5278
5027 AAGATGCAATTAATGCTTAAAGTAAACGAAGCGGATTAATCTTAAGCATGCACTTCAAGTG 5086

```
QY 5279 CAATATTGGAACAAATCACTTATGCTATTGTAACAAACGTAATAATTACACGGCGA 5338
DB 5087 AAATGCTGGAAGCCGAAACAAAACCTTGCTGAATTAAAACAACTCGCGATCAAAATGTTA 5146
QY 5339 ATGATCAAGATAGAAAAATATCATGTGACCTGCAAAAAGATTATACATGACCAATACGA 5398
DB 5147 ATCAAGCTACTCTTAAAGATGACATTGAAGTTCAATTCATATGACTTACATTATATTA 5206
QY 5399 TAGCAGTCGAGTTGAGAGAGCAAAAGAGCCTTGTCGAAGAGCTTCTGCAAGTACTA 5458
DB 5207 ACGATTACACATTCACACAGGTAAAGATGACGATACACACATTTATATGCTTATG 5266
QY 5459 CCTGAATATAGACAGTTCTTCTCATGTTGATCAAACTGATATTTGACAAAGATTAGAG 5518
DB 5267 CAGATCAGAAAGAAATATATATTTACAGCTGACATTAATGCAACACAGATGAAAAAC 5326
QY 5519 AAGAAATATATGAAATATAGAAAAAGCAAAATGTTATGTTCTAGCTGAAAAATACGAGTC 5578
DB 5327 AAGCAATTAAGCAAGTTGACCAAAATGTTCA--AACTGCAATTAGAAAAATTAAATG 5383
QY 5579 AAGTGTCACAAATGCGACAGTCTTTCCGAGCAAGTGCACAGCTGCAGTAGAGCTG 5638
DB 5384 GGTGTGATTAATGCTGACGTTGATGATGATTAACCAAGGTAAAGCAGCAATTGATGCTA 5443
QY 5639 GAGTAGCAGTTAATTAATTAACAAATATCTTCTGCACATTTAAAAATAGTACTCAA 5698
DB 5444 TTCAAGTAGATGCTACTGTTAACTTAAAGCGAACCAAGCTATTGAAGCTAAAGCAGAAG 5503
QY 5699 AT 5700
DB 5504 AT 5505
```

RESULT 30

```
AAA70212
ID AAA70212 standard; DNA; 1998 BP.
```

```
AC AAA70212;
```

```
DT 07-NOV-2000 (first entry)
```

```
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:345.
```

```
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
```

```
KM antimalarial; malaria; protozoasidae; infection; insecticide; ds.
```

```
XX Plasmodium falciparum.
```

```
XX WO200025728-A2.
```

```
XX 11-MAY-2000.
```

```
PF 05-NOV-1999; 99WO-US026796.
```

```
PR 05-NOV-1998; 98US-0107131P.
```

```
PA (HOPE/) HOFFMAN S.
```

```
PA (CARU/) CARUCI D.
```

```
PA (GARD/) GARDNER M.
```

```
PA (VENT/) VENTER J C.
```

```
PI Hoffman S, Carucci D, Gardner M, Venter JC;
```

```
DR MPI; 2000-365347/31.
```

```
XX Proteins encoded by chromosome 2 of the human malarial parasite,  
PT Plasmodium falciparum, useful as antimalarial vaccines and in the  
XX diagnosis of P. falciparum infection.
```

```
PS Disclosure; Page 537; 577pp; English.
```

```
XX The present invention describes proteins and their fragments (I) encoded
```

by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (i) nucleotide sequences (ii) encoding (i); and (2)
CC vaccines against P. falciparum infection comprising (i) or (ii); (i) and
CC (ii) are useful for the development of vaccines against P. falciparum
CC infection. (i) and polyclonal antisera or a monoclonal antibody raised to
CC immunogens comprising the sequences of (i), are useful in the detection
CC of infection with P. falciparum. Furthermore, (i) (especially when they
CC are rifins or secreted or membrane proteins) can aid the identification
CC of drugs to treat or prevent P. falciparum infection, or they can be used
CC to identify drug resistance in P. falciparum. Sequencing of the
CC Plasmodium chromosome 2 and the subsequent identification of proteins
CC encoded by it will help to expand our understanding of parasite biology,
CC a process hampered by the complexity of the parasitic life cycle, and
CC provide new targets for vaccine and drug development. Parasite resistance
CC to drugs and mosquito resistance to insecticides have led to a resurgence
CC of malaria in many parts of the world, and there is a pressing need for
CC vaccines and new drugs. AAA70078 to AAA70287 and AAB1814 to AAB18352
CC represent nucleotide and protein sequences given in the present
CC invention, but which are not specifically mentioned within the
CC specification

Sequence 1998 BP; 1127 A; 141 C; 338 G; 392 T; 0 U; 0 Other;

Query Match 0.7%; Score 68.8; DB 3; Length 1998;

Best Local Similarity 42.5%; Pred. No. 0.0031;
Matches 434; Conservative 0; Mismatches 577; Indels 9; Gaps 1;

```
QY 5130 TGATGAAACGGAACCTTATGATGATCTGAGTTGAAAGAGTAAGTCTTTCATGCT 5189
DB 816 TGATGAAAGATGAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 875
```

```
QY 5190 AGATGCAAAAGATCAAAACAAATTAATCAATTTGCCGAATGCAATGAGAGAAAGC 5249
DB 876 TGATGATTAAGAAAGAAACGATTAACACATTTGAAAGAAAGAAAGAAATTAATGGA 935
```

```
QY 5250 GGCTGAGTTGG-----AGCAACAGTTCCTCATCAATATTGGAAAAACATCAGT 5300
DB 936 AAGGAATTTTGTGATTAAGAAAGAAAGATGTAAGATTAAGATTAAGAAAGAAAG 995
```

```
QY 5301 TATAGCTATTGTAAGAAACAGTAATAATTACACCGGCAATATCAAGATAGAAAAATAT 5360
DB 996 TAAAGATATCTGGAAGAAAGAAAGAAAGTAAGATTAAGAAAGAAAGAAAGAAAG 1055
```

```
QY 5361 CAATGTGCTGCAAAAGATTATCTATATCAATCTATAGCAGTCGAGTTGGAGAGC 5420
DB 1056 GAAAGAAAAAGTAAGATTAAGAAAGAAAGAAAGTAAGATTAAGAAAGAAAGAAAG 1115
```

```
QY 5421 AAAAGAGCCTCTGCAAGGAGCTTGCAGTACTCTCACTTAATTAAGACAGTTCTTC 5480
DB 1116 TAAAGATATTGGAAGAAAGAAAGAAAGATTAAGATTAAGAAAGAAAGAAAGTAAG 1175
```

```
QY 5481 TCATGTGATCAAACTGATATTTGACAAAGATTAGAGAAAGAAATTAAGAAATAGGA 5540
DB 1176 TACTGCAAGGAAAGAAAGAAAGAAAGATTAAGATTAAGAAAGAAAGAAAGATATGA 1235
```

```
QY 5541 AAAGCAAAATGTTAATGTTTCTAGCTGAATTCAGAGTCAAGTGTCACAAAATGCCACGT 5600
DB 1236 AAAATTAAGATTAAGAAAGAAAGTAAGAAAGAAAGAAAGCGTTAAGAAAGAAAGAA 1295
```

```
QY 5601 GCTTCCGAGCAAGTGAACAAGCTGAGAGCTGAGAGCTGAGAGCTTAATTAATTAAC 5660
DB 1296 TGATTAACAGATTAATCAATGATGATTAATGATGATTAATGATTAATGAAAGAAATG 1355
```

```
QY 5661 ACAAATATCTTCTGCACATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5720
DB 1356 AAATGATGACGAAGAAAGATGAGAGATGAGAAACAAAGAAAGAAAGAAAGAAAG 1415
```

```
QY 5721 AAGCAATCTCATCTATTAAACAAATTTGAGATTTGAGAGTTGGAGCTGGAGG 5780
DB 1416 AAAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1475
```

```
QY 5781 AGCTGAGTGAAGAGTTCTGTGACGATGAATTAAGATTGTAATTAATTAATTAATTA 5840
```


pool of a type of animal, or individual sera with antibodies against the specific pathogen, tumour, allergen, tissue or host prone to auto-immunity. The hyperimmune serum-reactive antigens comprising any of the 62 sequences of 53-2261 amino acids fully defined in the specification, or their hyperimmune fragments are useful for the manufacture of a pharmaceutical preparation, particularly a vaccine against staphylococcal infections or colonisation against *S. aureus* or *S. epidermidis*. The preparation of antibodies is useful for the manufacture of a medicament for treating or preventing staphylococcal infections or colonisation against *S. aureus* or *S. epidermidis*. The antibody preparations may also be used for diagnostic and imaging purposes. Other conditions that can be treated include cancer, autoimmune diseases or infections caused by viral (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This polynucleotide sequence represents staphylococcal DNA relating to the method for identifying and producing pathogen specific antigens of the invention

Sequence 4911 BP: 2057 A; 938 C; 1030 G; 886 T; 0 U; 0 Other;

Query Match 0.7%; Score 68.2; DB 8; Length 4911;
Best Local Similarity 42.3%; Pred. No. 0.0054;
Matches 505; Conservative 0; Mismatches 683; Indels 7; Gaps 2;

```
Qy 5145 TTGAGTGTGATGAGTTGAGAGTAAGTTCTTCAATGTGATGCAAAAGATCA 5204
Db 93 TTGAGTGTGATGAGTTGAGAGTAAGTTCTTCAATGTGATGCAAAAGATCA 152
Qy 5205 AAAACAATTAATCAATGCGGAAATGCAATGAGGAAAGCGGTGAGTTGAGC 5264
Db 153 TAAACAATGCAAAAGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 212
Qy 5265 AACAGTCTCATCAAAATATTGAAACATCAGTTATAGTTATTTGTTAAACAGTGA 5324
Db 213 AACATTTAAGCAATCTTGTGCAAAATGAAATGAAATGAAATGAAATGAAATGAA 272
Qy 5325 AATTCAACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5384
Db 273 AACTTAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 332
Qy 5385 TATGACCAATCTATGAGAGTGTGAGAGTGTGAGAGCAAAAGAGCGCTGTGCAAGAGC 5444
Db 333 ATCAACAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 392
Qy 5445 TTGTGCAAGTACTTACCTTGAATTAAGACAGTTCTTCTCATGTTGATCAAACTGATATTGA 5504
Db 393 TGAAGAAACAGACAAAGTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
Qy 5505 CAAAGATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5564
Db 453 AACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 506
Qy 5565 TGAAGATGAGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5624
Db 507 ACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 566
Qy 5625 TGCAAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 5684
Db 567 AAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626
Qy 5685 AATATGATCTCAAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5744
Db 627 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 686
Qy 5745 AACCAATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 5804
Db 687 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 746
Qy 5805 AGTGAATTAAGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5864
Db 747 AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 806
Qy 5865 GGAATATGTCGAGTTATTAAGAGTGTGATGAGAGTGTGATGAGAGTGTGATGAGAGTGT 5924
```

```
Db 807 AGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 866
Qy 5925 GTCTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5984
Db 867 TAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
Qy 5985 TACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6044
Db 927 TGCTAAAGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 986
Qy 6045 TACTACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6104
Db 987 AGTATTAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1046
Qy 6105 CGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6164
Db 1047 TGAAGTACTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
Qy 6165 AGTACTCATCTTAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6224
Db 1107 AGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166
Qy 6225 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6283
Db 1167 TGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1226
Qy 6284 CTATATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6338
Db 1227 CCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1281
```

RESULT 33

ACA47636
ID ACA47636 standard; DNA; 4997 BP.

ACA47636;

19-UN-2003 (first entry)

Prokaryotic essential gene #29293.

Antisense; de; prokaryotic essential gene; cell proliferation;

drug design; gene.

Staphylococcus haemolyticus.

WO20027183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW,

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

P-PSDB; ABU43766.

New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

PS Claim 14; SEQ ID NO 3506; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX

Sequence 4997 BP; 2206 A; 790 C; 905 G; 1096 T; 0 U; 0 Other;

Query Match 0.7%; Score 68.2; DB 8; Length 4997;

Beet Local Similarity 41.0%; Pred. No. 0.0054;

Matches 933; Conservative 0; Mismatches 1318; Indels 24; Gaps 5;

Qy	5215	AATACAAATTCGCGAAATGCAAAATGAGAAAGCGCTGAGTGGAGCAACAGTTGCT	5274
Db	1969	AAAGCAGCAGCAAGAAAGCAATGATGATGACGCACTGCTCAAAAAAGAACGTAAT	2028
Qy	5275	CATACAAATATTGGAACCAATCACTTATAGCTATTGTAACCAAGTAATTAACAAG	5334
Db	2029	AATAGCTCAGATGCAACACAGAAAGAAAGATGACCAATAGCAAAAGTTGATCAGCT	2088
Qy	5335	GCGATGATCAAGATGAGAAATATCAATGATGCAAAAGATTAATCAATGCAAT	5394
Db	2089	GTAACAGCTGCAAAACAGCAATTACACAGCAACACAAATGATGAT--GTAGCCAG	2146
Qy	5395	ACTATGACAGTCGAGTTGAGAGCAAAAGAGCCTCTGTCGACGAGCTTCTGCAAGT	5454
Db	2147	AACAAATAGCGGTACTTCTGACTATTACTGATACAAACAG--AAGTTATCAAAAAACA	2205
Qy	5455	ACTACCTTGAATPAGACAGTTCTTCTCATGTTGATCAATGATATGACAAAGTTTA	5514
Db	2206	GCAGCAAGAAAAAGCAATTGATGATGAAGTCGTGTAAGAAAGCAATGATCTGTT	2265
Qy	5515	GAGGAGAAATATATGAAATTAAGGAAAGCAATGTTATGTTCTGCTGAAATACG	5574
Db	2266	GCCGATGCTACAGATGAGAGAAACAGGCAAGCTAAAGATTAAGTTGATGTAAGCTACC	2325
Qy	5575	AGTCAAGTGTGTCACAAATGCGCAGTGTCTTCGAGAGCAAGTGCACAGCTGAGTAGA	5634
Db	2326	AAAGCAAAAGCAGCAATGATGATCACTACTACAAACAGATGTAAGACAGCAAAATCA	2385
Qy	5635	GCTGAGTAGCAGTTATTAATTAATTAACAAATATCTTGCACATATTAATAATGTAAT	5694
Db	2386	TCTGTGTAAGCAACTATTGAAGAAATACAACTGATTAAGAAATTCGACGCAAAA	2445
Qy	5695	CAAAATGTAAGCAATGCTTGTGTAACAAATTCATTCATCTATTAATTAATGGA	5754
Db	2446	CAAGCAATAGATGATGTCACAAATGCTTAATAAGACAGCAATGATCAATATATGATGCA	2505

Qy	5755	ATTGAGCTGAGTGTGAGAGCTGAGAGCTGAGTGAACAGTTCTGTAGCAGTAATPAG	5814
Db	2506	ACTCAAGAGAGAAAGAGCTGAGCAAAAGTAATGATGCAAGCTGTCAAGCTTAAT	2565
Qy	5815	ATTGTAATTAATAGCATAGCAATTAATCATGCAAAATCATCTGCAAGGAAATGTC	5874
Db	2566	ACACCGATTGATCAAGCAACCAAGATGATGTCACAAAGAGAGACTTAGGCAAC	2625
Qy	5875	GGAATTTTACAGAGTCTGATGCGGTATGCTTAATTAATGAGAGAAACAGTCTGAGT	5934
Db	2626	AATTCATTTTCAATTAATTCAGCTGATATTTGTTAAAAAGAGAGCAAGAAAGCAAT	2685
Qy	5935	GCCCGTGCAGCAATGAGAGCTCAACAGTGTGATGAATTAACAGATCTTCAAAAG--	5992
Db	2686	GATATGACAGCAAGCTTAAGAAACAGAAATTTGATCAACCCCAACAGCTACCAAGAA	2745
Qy	5993	----CATATGTAAGATTTCTACAGTATGCTTAAGAGAAACAGATGATTAATTAAT	6048
Db	2746	GAAGAGAGCAGCAAGCAAGCTTAAGTATGCTGAAGTATTAAGTATTAATTAACGAA	2805
Qy	6049	ACTCAAGGCGCAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAATTAACGAA	6108
Db	2806	ACTCAAGCAATTTCAAAATTAATGATGATGATCAAAATTTCTGAACTGCTCAATA	2865
Qy	6109	GACTTATCAAAAAAGAAAAATTAAGTATTAATTAATTAATTAATTAATTAATTAAGCT	6168
Db	2866	GCAGATTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2925
Qy	6169	ACTCATACTTTAAATCTTTATTTGCAATGCGCTGCTGTTGAGACAGCCGAGTGCA	6228
Db	2926	GCAGCTAAGAAAAATGAAATGACCAATCCACACTACACAAAGAAAGAAAGATGCA	2985
Qy	6229	GGAATGTTAAT--CAACAGTTTATGAGAAACAGAGCTTGTGAAATTTCT	6285
Db	2986	GCAAGAGCAAAAGTATGAAAGATGCTTAACCAAAACAGCAATTTGATTCATCGACT	3045
Qy	6286	ATATTAATGCAAAACATTTATCTGTAAATCAAGAGATTAACAGAAATTCATGAGTA	6345
Db	3046	ACGAATGACGCGTATGATCAAGCAAAAGCTCTGTGTATCAATTAATTAATTAATCA	3105
Qy	6346	GTAGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	6405
Db	3106	CCAAATCAATTAAGAAAAATTAAGCAAGCAAGCTGTTGATCTGTTACAGCCCAA	3165
Qy	6406	ATTAAGAAATTAACAGACAGAGTTGAAAATCAATGTCATGTAAGTGAAGTTTGG	6465
Db	3166	AAGAACGCTATGATCAAGATGTAATCACTACAGAAAGAAAAAGATGCTTAAGCA	3225
Qy	6466	GAAAGCTGAATTAACAGCAGATTTCTAAGCAAGAAATTTCTTTTGGAGTCGAGTC	6525
Db	3226	AAAGTATGAAAGAGTAACTAAAGCAAACTTATGATCAAGCTATGATATGAT	3285
Qy	6526	GCAGCAGCCGCGTGAAGAGCCGAGTGGCAGAAACGTTTCCGTAATCAATTTGCAAGA	6585
Db	3286	GCTGTACAGCAACTCGAGATTAATGAATAATCTGTTAATAAATTTCAACTGAGACA	3345
Qy	6586	AAAGCAGAGTATGATGAGAAAGCAAGATTTTGTGTAAGAAAAAGCTGAGATTAACGA	6645
Db	3346	ATCAAAAAAGCTGATGAAACAGCAATCGATATGCTGTACAGCTTAAGAAAAAGCAGCA	3405
Qy	6646	AAAGCTATAGTTCTGTTGCAATTTGAAATGCGCAGTGGAGTGGCTGCAAAAGAGACT	6705
Db	3406	ATATATCAAGTTTCTGATGCTACAGAGAGAGAAAGATGAGGAAAGCTTAAGGTTGAC	3465
Qy	6706	GGAATGAGCAGCAGTGGCAGTTACCAAGATGAATCAACAGAGACAGAGAGTGA	6765
Db	3466	GAAAGATTAACAAAGCTAAGTCAAGCATGATTAAGCAGCAACAAATTAATGATGTTGAC	3525
Qy	6766	AATTTTAATTAATGATCTCGAACAAGTATGATGATGATGATGATGATGATGATGATG	6825
Db	3526	CAAGCAAAACGAAATGGAATCAATTAATTTCTTCAATTAAGAACGAG-----A	3573
Qy	6826	GATACGTAATCGGTTGACGCGAGCTGGAATTTCTTGACGCGAGATGATCTGAGTGT	6885

XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
 KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 OS Homo sapiens.
 PN W0200168912-A2.
 PD 20-SEP-2001.
 PF 15-MAR-2001; 2001WO-EP002955.
 PR 15-MAR-2000; 2000DE-01013847.
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 PA (EPIG-) EPIGENOMICS AG.
 PI Olek A, Piepenbrock C, Berlin K;
 DR WPI; 2001-602752/68.
 XX Fragments of chemically modified genes associated with tumor suppressor
 PT genes and oncogenes, useful in designing primers and probes for analyzing
 PT diseases associated with cytosine methylation state e.g. cancer.
 XX
 PS Claim 1: SEQ ID NO 409; 27pp; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and oncogenes
 CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
 CC 500 are missing from the sequence listing) sequences (Ss) and sequences
 CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
 CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
 CC probes for detecting the cytosine methylation state and/or single
 CC nucleotide polymorphisms and also to be used in an array for analysing
 CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The
 CC probes can also be used in a method for ascertaining genetic and/or
 CC epigenetic parameters for the diagnosis and/or therapy of existing
 CC diseases or the predisposition to specific diseases, by analysing
 CC cytosine methylations. The parameters may be compared to another set of
 CC genetic and/or epigenetic parameters, the differences serving as basis
 CC for diagnosis and/or prognosis events which are disadvantageous to
 CC patients. The present sequence is one of the 533 genomic sequences
 CC derived from tumour suppressor genes and oncogenes. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 7442 BP; 1655 A; 214 C; 1934 G; 3636 T; 0 U; 3 Other;

Query Match 0.7%; Score 67.6; DB 4; Length 7442;
 Best Local Similarity 45.4%; Pred. No. 0.0079;
 Matches 328; Conservative 0; Mismatches 384; Indels 10; Gaps 2;

QY 9012 TCTTGATGTAATTAATGATTAAGAAAAATGAGAGCAAAATCGAGAGCATGCTAT 9071
 DB 2679 TATTAAAAAATAAATACTAATAAAAAATAAATACTAATAAAAAATGAAATTT 2620
 QY 9072 TCTAGAACTACTGGA-----AACAAAGATATCAAGCATTTCAAGAGCAAAAGTAATA 9127
 DB 2619 AAAAAAATAAATACTAATAAAAAATAAATACTAATAAAAAATAAATACTAATA 2560
 QY 9128 TTCTTGAAAAAGAGACGCTGACAGTGCAGTCAATATGATGTACACATTTCCAATG 9187
 DB 2559 TAAAAAATAAATACTAATAAAAAATAAATACTAATAAAAAATAAATACTAATA 2500
 QY 9188 AGATGATATTTAAAAATTTGGCAAGAGCATGATGCTTCTCAATTATACCAAAAT 9247

DB 2499 TAAAAATAAAAATAAAATTCGAAAAAATAAATAAATAAATAAATAAATC 2440
 QY 9248 CAAAAAATAATATTACTTTAGCATCAAGTAGTAGTAATGTAATGTTGATGGGTGG 9307
 DB 2439 TAAAAAATAAATACTAATAAAAAATAAATACTAATAAAAAATAAATACTAATA 2380
 QY 9308 CTGAAGCAAGAGTGCGAGAGCCAAAGCGACGTTAGTGTAAAGATCAAAATAATAGAA 9367
 DB 2379 AAAAAAATAAATCAATATAAAAAATAAATAAATAAATAAATAAATAATCTAA 2320
 QY 9368 CTAAATAATGTTGATTAGCGAGAAATAATTAACAAGAGCAAAATCAATGTAATGCCG 9427
 DB 2319 AAAAAAATAAATACTAATAAAAAATAAATAAATAAATAAATAAATAAATACTAA 2260
 QY 9428 GATATGATAAAAATTAATATATATAGTAAGACAAATTCGAAGCTATTGCGAGTCCAAA 9487
 DB 2259 AAAAAAATAAATCTAATAAAAAATAAATAAATAAATAAATAAATAAATAAATACTAATA 2200
 QY 9488 GTCATGCTGACGCTGCTGGCAACTGCCACTATTGAAAAAATAAGTAAATTT---- 9543
 DB 2199 AAAAAAATAAATCACTAATAAAAAATAAATAAATAAATAAATAAATAAATAAATA 2140
 QY 9544 --AATAATGCGATCCGAAATTTAAATAATACTGGCAAGATTGGAAGGAAAGCTAATA 9601
 DB 2139 AAAAAAATAAATACTAATAAAAAATAAATAAATAAATAAATAAATAAATACTAATA 2080
 QY 9602 AAAAAAGCTCGGTAGATCTAATCAGGTAGACTGTATACGATTAATATACATGCAATT 9661
 DB 2079 AAAAAAATAAATACTAATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2020
 QY 9662 CTCTGAAAAAGCATACAAAAAATTGACATATCAATCAAGAGAGAGAAAAAGGAAA 9721
 DB 2019 AATCTAAAAAACACTAATAATAAATAAAGCACTCTAATAAAAAATAAATAAATAAATA 1960
 QY 9722 AA 9723
 DB 1959 AA 1958

RESULT 36
 AAL53680/C
 ID AAL53680 standard; DNA; 15507 BP.
 AC AAL53680;
 XX
 AC 07-FEB-2003 (first entry)
 XX
 DT Genomic DNA sequence part 1 of the human ch10-ARPR protein.
 XX
 DB Cytostatic; chromosome 10 arginine-rich protein; ch10-ARPR; ARPR; cancer;
 KW Cytostatic; chromosome 10 arginine-rich protein; ch10-ARPR; ARPR; cancer;
 KW gene therapy; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH 5'UTR 927..962
 FT /*tag= a
 FT 963..13435
 FT CDS
 FT /**tag= b
 FT /product= "Part of the ch10-ARPR human protein"
 FT /note= "No stop codon. Coding sequence contains 2
 FT introns"
 FT 963..1077
 FT /**tag= c
 FT /number= 1
 FT 1078..10639
 FT /**tag= d
 FT /number= 1
 FT 10640..10767
 FT /**tag= c
 FT /number= 2
 FT 10768..13294
 FT intron

```
FT      /*tag= d
FT      /number= 2
FT      exon      13295..13435
FT      /tag= c
FT      /number= 3
PN      WO200279246-A2.
PD      10-OCT-2002.
PF      26-MAR-2002; 2002WO-EP003395.
PR      30-MAR-2001; 2001US-0280673P.
PR      24-MAY-2001; 2001US-0293453P.
PA      (GENE-) GENEPROT INC.
PI      Bouguetelret L, Niknejad A, Balroch A;
DR      WPI; 2003-040654/03.
XX      New isolated chromosome 10 arginine-rich protein related polypeptides,
XX      useful for detecting and/or monitoring and treating conditions involving
XX      aberrant expression of ARPR or uncontrolled growth of tissues, such as
XX      cancer.
XX      Claim 8; Page 55-65; 82pp; English.
XX      The invention relates to isolated chromosome 10 arginine-rich protein
XX      related polypeptides (ch10-ARPR), or their fragments. The polypeptides,
XX      nucleic acids and antibodies are useful for detecting and/or monitoring
XX      and treating conditions involving aberrant expression of ARPR or
XX      uncontrolled growth of tissues, such as cancer. The polypeptides are
XX      useful as hybridisation probes, in chromosome and gene mapping, for the
XX      generation of antisense RNA or DNA and in tissue or cell typing. The
XX      methods are useful for detecting and measuring quantities of ARPR in
XX      tissues and biological fluids. The host cells are useful for replicating
XX      ARPR transcripts or expressing the ARPR proteins or polypeptides. The
XX      polynucleotides of the invention can be used to treat disorders by gene
XX      therapy. This sequence represents the genomic DNA sequence, part 1, of
XX      the human ch10-ARPR protein of the invention
XX      Sequence 15507 BP; 4046 A; 3576 C; 3266 G; 4618 T; 0 U; 1 Other;
SQ
Query Match      0.7%; Score 66.2; DB 8; Length 15507;
Best Local Similarity 50.8%; Pred. No. 0.018;
Matches 158; Conservative 0; Mismatches 153; Indels 0; Gaps 0;
QY      7105 GTAACGAGAAAAATGTAATGTAACGCAAAAGAGAAAAAATTTAAGCAACAGCAGCA 7164
DB      12860 GAAAGAAAGAAAGAAATGAAATGAAAGAAAGAAAGAAAGAAAGCAATGCAGCAGCAGCA 12801
QY      7165 AATGCAGGAATGGAGAGCAGCAATCGAGCCAAATGCTTGGTAAATTTTGAACA 7224
DB      12800 GCAGCGCAGCAGCAAGACAGAGAAAGAAAGAAAGAAAGAAAGAAAGCAAGCA 12741
QY      7225 GCTGTGAAGATTAAGAAAAATTTGTAAGAAAGAAAGAAAGAAAGAAAGCAAGCA 7284
DB      12740 GCAGCGCAGCAGCAAGCAAGCAGCAAGAAAGAAAGAAAGAAAGCAAGCAAGCA 12681
QY      7285 GAAAGTAAACAAAGAAACAGTAAATTAATGATGCTACGAAAAAATCTTACATCA 7344
DB      12680 GCAGCGCAGCAAGAAACATTAAGAGCACTTAAGAAAGCAAGCAAGCAAGCAAGCA 12621
QY      7345 GCAGGATATTTCTTACAGAGATCTTGTAAACCGATAGAGAGATTAAGCAAGCA 7404
DB      12620 GCAGAGCAGCAGCAAGCAAGCAAGCAAGTGAAGGATTCAGAAAGCAGCAAGCAAGCA 12561
QY      7405 GGAATTTAAAGC 7415
DB      12560 GCAGCAGAAAGC 12550
```

```
RESULT 37
ADQ62833
ID      ADQ62833 standard; RNA; 1000 BP.
XX      AC      ADQ62833;
XX      DT      07-OCT-2004 (first entry)
XX      DE      Homopoly-A contaminant for RNaseH activity assay.
XX      KW      ss; nuclease-mediated cleavage; target nucleic acid; RNaseH activity;
XX      KW      fluorophore; fluorescence quencher; fluorescent signal; HIV;
XX      KW      reverse transcriptase.
XX      OS      Synthetic.
XX      PN      WO2004059012-A1.
XX      PD      15-JUL-2004.
XX      PF      22-DEC-2003; 2003WO-US040879.
XX      PR      23-DEC-2002; 2002US-0436125P.
XX      PA      (AMHP ) WYETH.
XX      PI      Olson MW, O'Connell JF;
XX      DR      WPI; 2004-543471/52.
XX      PT      Detecting a nuclease-mediated cleavage of a target nucleic acid, useful
XX      PT      for detecting and monitoring RNase H activity, comprises hybridizing a
XX      PT      target nucleic acid to a fluorescently labeled oligonucleotide probe.
XX      Example 1; SEQ ID NO 4; 61pp; English.
XX      The invention relates to a method of detecting a nuclease-mediated
XX      cleavage of a target nucleic acid or measuring a RNaseH activity of an
XX      agent by hybridizing a target nucleic acid to a fluorescently labelled
XX      oligonucleotide probe complementary to the target nucleic acid and
XX      containing a fluorophore at one terminus and a quenching group at the
XX      other terminus and contacting the probe-target hybrid with an agent
XX      having nuclease activity. When the oligonucleotide probe is unhybridised
XX      to the target nucleic acid, the probe adopts a conformation that places
XX      the fluorophore and quencher in such proximity that the quencher quenches
XX      the fluorescent signal of the fluorophore. Formation of the probe-target
XX      hybrid causes sufficient separation of the fluorophore and quencher to
XX      reduce quenching of the fluorescent signal of the fluorophore followed by
XX      contacting the probe-target hybrid with an agent having nuclease activity
XX      to selectively cleave the target nucleic acid and thus release the intact
XX      probe. Detection of the release of the probe is by measuring a decrease
XX      in the fluorescent signal of the fluorophore as compared to the signal of
XX      the probe-target hybrid. The methods are useful for detecting and
XX      monitoring RNaseH activity in real time. The methods are also useful for
XX      monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
XX      the invention, the activity of the RNaseH from E. coli or HIV reverse
XX      transcriptase is measure by the method of the invention. In order to test
XX      the specificity of the RNaseH cleavage of the target-probe complex, a DNA
XX      or RNA contaminant is added to the assay reaction. This sequence
XX      corresponds to the single stranded homopolymeric polyA RNA contaminant
XX      for testing this activity.
SQ      Sequence 1000 BP; 1000 A; 0 C; 0 G; 0 T; 0 U; 0 Other;
Query Match      0.7%; Score 65.8; DB 12; Length 1000;
Best Local Similarity 41.8%; Pred. No. 0.0098;
Matches 403; Conservative 0; Mismatches 562; Indels 0; Gaps 0;
QY      8759 AAATTGAGATTAATTAGAAAGATTGCAAGATTAATGATTTGTAGAGTCAACGAG 8818
DB      1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60
QY      8819 ACGGAACCAAGAGAGTCTTGTGCGAAGAAAGATGTAATTTCTGTGAAAAATACATTTCAG 8878
```

```

Db      61  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 120
Qy      8879  GGGAAACAAATCATCATTTGAAGATAAGCAGATTGTGGACCGGAATGTAAATG 8938
Db      121  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180
Qy      8939  TAGATGCTTGAATGAATCTGATGATCTACAGAAAGAAAGTGCGCTATGTGGAA 8998
Db      181  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 240
Qy      8999  TTGGATTGGAATGTGATGTAATATGTGATTAGAAAAATGTAGAGCCAAATCG 9058
Db      241  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 300
Qy      9059  GAAAGCATGCTATTGTAGAACTACTGGAAACAGAAATATCAAGCTTTCAAGACAA 9118
Db      301  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 360
Qy      9119  AAGTAATATCTTGGAAAGAGAGACGCTGACAGCTGACAGTCAATATCGAATGTACACA 9178
Db      361  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 420
Qy      9179  TTCCAAAGAGATGATTTTAAATTTGGCAAGCAGTATGCTCTCAATTAAATGA 9238
Db      421  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 480
Qy      9239  CCAAAATTTCAAAATATATATTACTTAGCATCAAGTAGTGAATGGAATGTC 9298
Db      481  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 540
Qy      9299  ATGGGTGGCTGAGCAAGGTGACAGAGCCAAAGCACTGTTAGTAAAGATCAAA 9358
Db      541  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 600
Qy      9359  TAAATGAATTAATATGTTGATTTAGCAGGAAATTTAAACAGAGGAAACATCAATG 9418
Db      601  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 660
Qy      9419  TATATGCGGATATGATATAATTTATATATATATAGTAAAGCAAAATCTAAGCTATTGGG 9478
Db      661  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 720
Qy      9479  ATGCCAAAGTCACTGCTGCTGCGCACTGCCACTATTGAAAAAATGAAGTAA 9538
Db      721  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 780
Qy      9539  AATTTAATATGCGATCCGAGATTTAAATATATCTGCGCAAGATTGGAAGGAAAGCTA 9598
Db      781  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 840
Qy      9599  ATAAACAAACGTCGGTAGATCTAATCAAGTAGACTGTATACGATTAATATCATGGC 9658
Db      841  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 900
Qy      9659  ATTCTTCTGAAAAAGCATACAAAAAATTGACATATCAATCAAGAGAGAAAAAGGGA 9718
Db      901  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 960
Qy      9719  AAAAAA 9723
Db      961  AAAAAA 965

```

RESULT 38
ADG62832/c
ADG62832 standard; RNA; 1000 BP.

ADG62832;

07-OCT-2004 (first entry)

Homopoly-U contaminant for RNaseH activity assay.

```

XX      ss: nuclease-mediated cleavage; target nucleic acid; RNaseH activity;
KW      fluorophore; fluorescence quencher; fluorescent signal; HIV;
KW      reverse transcriptase.
XX      Synthetic.
XX      MO2004059012-A1.
XX      15-JUL-2004.
XX      22-DEC-2003; 2003WO-US040879.
XX      23-DEC-2002; 2002US-0436125P.
XX      (AMHP ) WYETH.
XX      Olson MW, O'Connell JF,
XX      WPI, 2004-543471/52.
XX      Detecting a nuclease-mediated cleavage of a target nucleic acid, useful
XX      for detecting and monitoring RNase H activity, comprises hybridizing a
XX      target nucleic acid to a fluorescently labeled oligonucleotide probe.
XX      Example 1; SEQ ID NO 3; 61bp; English.
XX      The invention relates to a method of detecting a nuclease-mediated
XX      cleavage of a target nucleic acid or measuring a RNaseH activity of an
XX      agent by hybridizing a target nucleic acid to a fluorescently labeled
XX      oligonucleotide probe complementary to the target nucleic acid and
XX      contacting a fluorophore at one terminus and a quenching group at the
XX      other terminus and contacting the probe-target hybrid with an agent
XX      having nuclease activity. When the oligonucleotide probe is unhybridized
XX      to the target nucleic acid, the probe adopts a conformation that places
XX      the fluorophore and quencher in such proximity that the quencher quenches
XX      the fluorescent signal of the fluorophore. Formation of the probe-target
XX      hybrid causes sufficient separation of the fluorophore and quencher to
XX      reduce quenching of the fluorescent signal of the fluorophore followed by
XX      contacting the probe-target hybrid with an agent having nuclease activity
XX      to selectively cleave the target nucleic acid and thus release the intact
XX      probe. Detection of the release of the probe is by measuring a decrease
XX      in the fluorescent signal of the fluorophore as compared to the signal of
XX      the probe-target hybrid. The methods are useful for detecting and
XX      monitoring RNaseH activity in real time. The methods are also useful for
XX      monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
XX      the invention, the activity of the RNaseH from E. coli or HIV reverse
XX      transcriptase is measure by the method of the invention. In order to test
XX      the specificity of the RNaseH cleavage of the target-probe complex, a DNA
XX      or RNA contaminant is added to the assay reaction. This sequence
XX      corresponds to the single stranded homopolymeric polyU RNA contaminant
XX      for testing this activity.
XX      Sequence 1000 BP; 0 A; 0 C; 0 G; 0 T; 1000 U; 0 Other;
XX
Qy      Query Match 0.7%; Score 65.8; DB 12; Length 1000;
Db      Best Local Similarity 41.8%; Pred. No. 0.0098;
Qy      Matches 403; Conservative 0; Mismatches 562; Indels 0; Gaps 0;
Db      8759  AAATTGAGATTAATTAAGAAACGATTGCAAGAGATAATACGATTGTAAAGTCAACGAG 8818
Db      1000  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 941
Qy      8819  ACGGAACCAAGAGGCTCTTGCGGAAGAGATGATTTCTGTGAAAATATCAATTTGAG 8878
Db      940  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 881
Qy      8879  GGGAAACAAATCATCATTTGAAGATAAGCAGATTGTGGACCGGAATGTAAATG 8938
Db      880  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821
Qy      8939  TAGATGCTTGAATGAATCTGATGATCTACAGAGAAAGAGTGCGCTATGTGGAA 8998

```

```

Db      820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 761
Qy      8999 TTGGTATTGGAATTTGATGTAATTAATGATTAAGAAAAATGTAGACCAAAATCG 9058
Db      760 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 701
Qy      9059 GAAGACATGCTATTGTAGAACTACTGGAACAAGAAATATCAGCATTACAGAGCA 9118
Db      700 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 641
Qy      9119 AAGTAATATTTCTTGAAAAAGAGACGCTGACAGCTGACCTGCAATATCGAATGTACA 9178
Db      640 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 581
Qy      9179 TTTCCATGAGATGATATTAAAAATTTGGCAAGCAGTATCTTCTCATTAATTA 9238
Db      580 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 521
Qy      9239 CCAAAATTCAAAAAATTAATTAATCTTAGCTCAAGTATGATGCAATGCAATGTTTC 9298
Db      520 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 461
Qy      9299 ATGGGATGGCTGACGACAGAGGTGCGAGCCAAAGCGACAGTATGTAAGAAATCAA 9358
Db      460 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
Qy      9359 TAAATAGAACTAATAATGTTGATTAGCAGAAAAATTAAACAGAGGAAACATCATG 9418
Db      400 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 341
Qy      9419 TATATCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9478
Db      340 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 281
Qy      9479 ATGCCAAAGTCATGCTGACGCTGCTTCCGCAATCTCCATTTTGAATAAATGAATGA 9538
Db      280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 221
Qy      9539 AATTAATATGCGATCCAGAAATTTAAATTAATCTGCAAGATTGAGAGGAAGCTA 9598
Db      220 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 161
Qy      9599 ATAAAAAAACGTGCGATGATCTAATCAGGTAGACTGTATACGTAATATATCATGCG 9658
Db      160 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 101
Qy      9659 ATCTCTCTGAAAAAGCATCAAAAAATTGACATATCAATCAAGAGAGGAAAAAGGA 9718
Db      100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 41
Qy      9719 AAAAA 9723
Db      40 AAAAA 36

```

RESULT 39

AAA70105
ID AAA70105 standard; DNA; 5940 BP.

AAA70105;

07-NOV-2000 (first entry)

Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:238.

Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;

antimalarial; malaria; protozoicide; infection; insecticide; ds.

Plasmodium falciparum.

W0200025728-A2.

11-MAY-2000.

```

XX      05-NOV-1999; 99WO-US026796.
PF      05-NOV-1998; 98US-0107131P.
PR      (HOFF/) HOFFMAN S.
XX      (CARU/) CARUCCI D.
PA      (GARD/) GARDNER M.
PA      (VENT/) VENTER J C.
XX      Hoffman S, Carucci D, Gardner M, Venter JC;
PI      WPI; 2000-365347/31.
DR      Proteins encoded by chromosome 2 of the human malarial parasite,
XX      Plasmodium falciparum, useful as antimalarial vaccines and in the
PT      diagnosis of P.falciparum infection.
PS      Disclosure; Page 460-462; 577pp; English.
XX      The present invention describes proteins and their fragments (i) encoded
CC      by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC      Also described are: (1) nucleotide sequences (ii) encoding (i); and (2)
CC      vaccines against P. falciparum infection comprising (i) or (ii). (i) and
CC      (ii) are useful for the development of vaccines against P. falciparum
CC      infection. (i) and polyclonal antisera or a monoclonal antibody raised to
CC      immunogens comprising the sequences of (i), are useful in the detection
CC      of infection with P. falciparum. Furthermore, (i) (especially when they
CC      are refined or secreted or membrane proteins) can aid the identification
CC      of drugs to treat or prevent P. falciparum infection, or they can be used
CC      to identify drug resistance in P. falciparum. Sequencing of the
CC      Plasmodium chromosome 2 and the subsequent identification of proteins
CC      encoded by it will help to expand our understanding of parasite biology,
CC      a process hampered by the complexity of the parasitic life cycle, and
CC      provide new targets for vaccine and drug development. Parasite resistance
CC      to drugs and mosquito resistance to insecticides have led to a resurgence
CC      of malaria in many parts of the world, and there is a pressing need for
CC      vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
CC      represent nucleotide and protein sequences given in the present
CC      invention, but which are not specifically mentioned within the
XX      specification
SQ      Sequence 5940 BP; 3106 A; 343 C; 879 G; 1612 T; 0 U; 0 Other;
Query Match 0.7%; Score 65.4; DB 3; Length 5940;
Best Local Similarity 44.7%; Pred. No. 0.02;
Matches 472; Conservative 0; Mismatches 571; Indels 14; Gaps 5;
Qy      10 ATCAAAATTAAGCTTCAGAGGACAGGAAAGGATATGATTCGTAATAAAGTTTAAG 69
Db      136 ATAGAAACTAAAGGAAAAAGTAAATTAATAAATGATAATTCCTTATAGATGATAT 195
Qy      70 ATTTGGATTTGTGATTAACACTATGACGGTGGGCTATGATCAATCACCGGACT 129
Db      196 ATTAATGAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 255
Qy      130 GAGAAATTTGGAACAAATATAGAAAAAGATAATGTTTATGACATTCTCAACAAG 189
Db      256 AATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 315
Qy      190 ATTCAAGGGGAGACGCTTTTAACGTTTAAATAGATTGCTTTAAGAAATAATATA 249
Db      316 AATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 374
Qy      250 GCAATCTATATTTTGGGGAAGAAGATGTAAGGGGTAAATATCTTTTAACTTTGTC 309
Db      375 TAAAGCAATGAGCTGAGAAACACAGCTTAAGATATCAATTAAGTCATTAGTCGTGC 434
Qy      310 AATGAAAAATTAAGATAGATGATTAATCAACGAATTCGAGAAATAATTAATTTGAGGA 369
Db      435 GATATAAATTTGAAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 494
Qy      370 AATTTATATTTCTTAAGCTCGAAGGATGCGAGTAGGAAAAAATGCAATGATGCT 429

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2005, 03:43:28 ; Search time 22519 Seconds
(without alignments)
20207.432 Million cell updates/sec

Title: US-10-647-057-8
Perfect score: 9726
Sequence: 1 atgagcgccatcaataataa.....gagaaaaagggaataatga 9726

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsat1:*
10: gb_gsat2:*
11: gb_gsat3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	104	1.1	1566	10	CG57757 P053-1-D0
C 2	100.8	1.0	1811	10	CG573732 P048-4-G0
C 3	100.8	1.0	1917	10	CG573732 P048-4-G0
C 4	100	1.0	1452	10	CG573732 P048-4-G0
C 5	99.8	1.0	1843	10	CG573732 P048-4-G0
C 6	97.8	1.0	1491	10	CG573732 P048-4-G0
C 7	97.4	1.0	1434	10	CG573732 P048-4-G0
C 8	97.2	1.0	1434	10	CG573732 P048-4-G0
C 9	96.8	1.0	1519	10	CG573732 P048-4-G0
C 10	96.6	1.0	1691	8	DR124185 49179719
C 11	96.2	1.0	2001	10	CG573732 P048-4-G0
C 12	95.4	1.0	1547	10	CG573732 P048-4-G0
C 13	95.2	1.0	1784	10	CG573732 P048-4-G0
C 14	95	1.0	1721	8	DR145459 49056454
C 15	93.8	1.0	1482	10	CG573732 P048-4-G0
C 16	93.2	1.0	1416	10	CG573732 P048-4-G0
C 17	93.2	1.0	1531	10	CG573732 P048-4-G0
C 18	93	1.0	1594	10	CG573732 P048-4-G0
C 19	92.4	1.0	1654	10	CG573732 P048-4-G0
C 20	91.8	0.9	1256	10	CG573732 P048-4-G0
C 21	91.8	0.9	1485	10	CG573732 P048-4-G0
C 22	91.8	0.9	1808	8	DR125630 49056520

C 23	91.6	0.9	1528	10	CG573854 P049-1-D0
C 24	91.4	0.9	1453	10	AJ591978 Arabidops
C 25	91.4	0.9	1829	10	CG109409 ISB1-51H2
C 26	91.2	0.9	1981	10	CG108200 CH216-165
C 27	90.6	0.9	1507	10	AG346189 Mus muscu
C 28	90.2	0.9	1401	10	AG289820 Mus muscu
C 29	90.2	0.9	1407	10	AJ592026 Arabidops
C 30	90.2	0.9	1618	10	CG107236 CH216-25H
C 31	89.6	0.9	1824	10	CG1081234 CH216-161
C 32	89.2	0.9	1932	8	DR141720 49300102
C 33	88.6	0.9	1434	9	CG187638 CH261-98P
C 34	87.8	0.9	2071	8	DR147304 49027161
C 35	87.6	0.9	1457	10	CG1082658 CH216-169
C 36	87.4	0.9	1428	10	CG1081978 CH216-165
C 37	87.2	0.9	1139	1	AJ927291 AJ927291
C 38	87.2	0.9	1512	10	CG113943 ISB1-55B2
C 39	87.2	0.9	1783	8	DR141924 49190649
C 40	87	0.9	1650	10	CG1041163 CH216-51P
C 41	86.6	0.9	1438	10	CG1078561 CH216-151
C 42	86.6	0.9	1816	10	CG118629 ISB1-72B1
C 43	86.4	0.9	1314	10	CG1077082 CH216-143
C 44	86.4	0.9	1616	10	CG1081995 CH216-165
C 45	86.4	0.9	1753	8	DR131208 49217594

ALIGNMENTS

RESULT 1
CG57757/c
LOCUS P053-1-D07.zc Ppa EcORI BAC library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
ACCESSION CG57757
VERSION CG57757.1 GI:37986636
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE Srinivasan,J., Stutz,W., Jessee,T., Wieggers-Perreolte,L., Jansen,K.,
1 (bases 1 to 1566)
Buntjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
JOURNAL PUBLISHED
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES
source
1..1566
Location/Qualifiers
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcORI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcORI and cloning into the BAC
vector."
ORIGIN
Query Match 1.1%; Score 104; DB 10; Length 1566;
Best Local Similarity 37.0%; Pred. No. 4.9e-10;
Matches 482; Conservative 0; Mismatches 820; Indels 0; Gaps 0;
QY 8422 AACGAGATTATGAAATATATATACCAAAAAATATATTCGAGAGTCAATGCTCTT 8481

[illegible][illegible]

RESULT 2	CG753732/c	LOCUS	DEFINITION
	CG753732	1811 bp	DNA
	P048+G03.YA	BAC library	linear
	genomic survey sequence.		Q58 24-OCT-2003
			<i>Pristionchus pacificus</i> genomic.

ACCESSION	CG753732
VERSION	CG753732.1
GI	GI:37978509

KEYWORDS GSS.
SOURCE Pristionchus pacificus

ORGANISM *Priestionchus pacificus*
Eukaryota; Metazoa; Nematoda

Neodiplogasteridae; Pr
1 (bases 1 to 1811)
REFERENCE

AUTHORS
Srinivasan, J., Sinz, W.,
Buntjer, J., van der Meulen, J.

TITLE An integrated physical
pacificus

JOURNAL
PUBMED
MOL. Genet. Genomics 26
12884007

COMMENT:
Contact: Sommer RJ
Evolutionary Biology

Max-Planck-Institute for
Spemannstr. 37-39, Tuebingen

Tel: 00497071601371
Fax: 00497071601498

Email: fall.sommer@ueh.de
Class: BAC ends.

FEATURES	Location/Qual
Source	1. .1811

```
/organism="P. falciparum"
/mol_type="genomic"
/submitter="NCBI"

```

```

/strain="CALL
/db_xref="taxo

```

```
/clone_lib="perl"
/notes="The lib
```

the genomic DNA vector."

ORIGIN

Query Match	1.06;
Best Local Similarity	34.2%;
Watchdog ECT	Conservation

Marches 56 /; ConberValive 0;

8068 AAAG1IAG1ACAAA1LCGGH

DB 1/31 AAAAAAAAAAAAAAAAAA

8128 ACACCAIAAAAGGAAAGTAACT
1501 TTTTATTTTATTTTATTTTATTTT

DB 1691 AAAAAAAAAAAAAAAAAA

8188 GTTGAGTCACTAAGGCAGAT
| | ||
1071 TTTTATTGTTTAAATTAATTA

DB 1631 AANNNAANNAANNAANNA

CN 0348 GGNNTTCTTCAGACAT

8248 GGAAATTCACAGACAAAT
|||
1571 AAAAAAAAAATTTT

DB 13 / 1 AAAAAAAAAAAAAAAAAAAAAAAAAA

[illegible]


```

QY 8481 TGCCTTAATGATCAAGAAATGAGCGAATATAGAACTTTTACCGCTTACCGGTGTGCA 8540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 387 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 446
QY 8541 TGCACAGGACCAACAAAGCATTTTACGAGTCAACCAAGTTTAACTTCTACACTGTAA 8600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 447 AAAAAACAAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 506
QY 8601 TGCAGAAACGTATCTCAACTTCGTCAAAAGCTTTGCTTAAATATGAAATTTATGAAA 8660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 507 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 566
QY 8661 TGTAAAGGACCTGAGAGACCTTGTGCGAGCGAAACAGCAGCCCTTGAATATATAC 8720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 567 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 626
QY 8721 AAAGACTACTACAGAGCACTTGTTCAGAAATTTGGGAAATTTGAGATTAATTGAAAC 8780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 627 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 686
QY 8781 GATTGCAAGATTAATACGATTTGTAAAGTCAACGAGAGCGAAACCAAGAGAGTCTTGT 8840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 687 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAG 746
QY 8841 CGAAAGAAATGGTATTTCTGTGAAAAATCAATTTCAAGGAGAAACAAATTCATCATTA 8900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 747 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 806
QY 8901 AGATAAGCGAATTTGTTGGAACGGAAGTGTAAATGTAGATGCTTTGAATGAACCTTA 8960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 807 AAAAAACGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 866
QY 8961 TGTAGATCTACAGAAAGAAAGTGTGTGCTATGTGTGAATTTGTATTTGAAATGTTGATG 9020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 867 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 926
QY 9021 AAATATGTGATTAAGAAAAATTTAGAGCCAAATCGAGAGCATGCTATTTGTAGAAC 9080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 927 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 986
QY 9081 TACTGAAACAAAGAAATATCAACATTTTCAAGAGCAAAAGTAAATTTCTTGGAAAAAG 9140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 987 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1046
QY 9141 AGACGTGCACTGCTGCACTGCAATATCGAATGTACATTTCCAAATGATGATATTTAA 9200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1047 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1106
QY 9201 AAATTTGCAAGCACTATGATCTTCTCAATTAATACCAAAATTTCAAAATATATAT 9260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1107 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1166
QY 9261 TACTTAGATCAAGTATGTAATCGAATGTGAATGTTCAATGGGTGCTGAAGCAAGAG 9320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1167 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1221
QY 9321 TGCAGAGCCAAAGCAGCACTTGTATTAAGATCAATTAATTAAGAACTTAATATGTTGA 9380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1222 GAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1281
QY 9381 TTTTACAGAGAAATTTAAACAGAGGAAACATCAATGTATATGCGGATATGATTAATA 9440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1282 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1341
QY 9441 TTATATATATAGTAAGCAAAATTTCTAAGCTATTTGGGATGCGCAAAAGTCACTCTCAGC 9500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1342 AAGCAAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1401
QY 9501 TCGCTGCGCACTGCACTATTTGAAAAATTAAGAAATTTTAATATATGCGATCCGAGA 9560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1402 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1461
QY 9561 ATTTAAAAATTAATCTGCGCAAGATTGAGAGGAAAGCTAATTAATAAAACCTCGGTAGATC 9620

```

```

DB 1462 AAAAAAAAAAAAAAAAAAGAAACAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1521
QY 9621 TAATCAGGTAGACTGATATACGATTAATATATCATGCTCTTCTGAAAAAGCATACA 9680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1522 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1581
QY 9681 AAATTTGACATATCAATCAAGAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAA 9723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1582 AAAAAACAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1624

RESULT 4
CL125503      1452 bp      DNA      linear      GSS 05-JAN-2004
LOCUS        ISB1-87B20.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-87B20,
DEFINITION   genomic survey sequence.
ACCESSION    CL125503
VERSION      CL125503.1 GI:40619138
KEYWORDS     GSS.
SOURCE       Xenopus tropicalis (western clawed frog)
ORGANISM     Xenopus tropicalis
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE    1 (bases 1 to 1452)
              Kremetzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
              Mardis,E. and Wilson,R.
              A physical map of the xenopus tropicalis genome
              Unpublished (2003)
              Contact: Richard K Wilson
              Genome Sequencing Center
              Washington University School of Medicine
              Email: submissions@wustl.edu
              Insert length: 75000 Std Error: 0.00
              Seq primer: T7 TAATCAGACTCACTATAGGG
              Class: BAC ends
              High quality sequence start: 672
              High quality sequence stop: 905.
FEATURES             location/Qualifiers
     source           1..1452
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="ISB1-87B20"
                     /clone_1ib="ISB1"
                     /note="Vector: pBelobAC11, ISB-1 Xenopus tropicalis BAC
                     library Segment 1"

ORIGIN
Query Match      1.0%; Score 100; DB 10; Length 1452;
Best Local Similarity 41.3%; Pred. No. 3.1e-09;
Matches 558; Conservative 0; Mismatches 792; Indels 1; Gaps 1;

QY 8374 AATCTTTACAGCAAAAGTAATCTGAATCGTAGTTGTTTACGAAGCAAGATTAT 8433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 23 AAAAAACCAACAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 82
QY 8434 GAAATTAATGATTAACACAAATTAATATTTTGAAGTCAATGCTCTTCTTAATGAT 8493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 83 AAAAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 142
QY 8494 ACAAGAAATGAAGCAATTAATCTTTAGCGGTAGCCGTGTGCATGCACAGAGACA 8553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 143 AAAAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 202
QY 8554 AACCAAGATTTACAGATCAACAGTTTAACTTCTACAACTGTAAATGAGAGAAACGTA 8613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 203 AAAAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 262
QY 8614 TCTCAACTTCGTGCAAAAGCTTGGCTAAATTAATTAATTAATTAATTAATTAATTA 8673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 263 CTAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 322

```

```

QY 8674 GAGAGAGCTTACTCGAGCGGAACAGACGCCGTTGAAATTTATACAAAGACTACTACA 8733
Db 323 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 382
QY 8734 GGAGCATGGTGTTCAGGAAATTTGGAAATTTGAGATTAATTTAGAAACGATTCGAAAGAT 8793
Db 383 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 442
QY 8794 AATACATTGTAAAGCTCAACGAGACGGAACCAAGAGAGTCTTGTGGAAGAAATGCT 8853
Db 443 AAAAAAAAAAAAAAAAAAAGAAAGAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 502
QY 8854 ATTTCGTGAAATTAATCAATTTGAGGGGAAACAAATTCATCCATTGAAATTAAGCCAGA 8913
Db 503 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 562
QY 8914 ATTGTTGAAACCGAAGTGTAAATGTAGATGCTTTGAAATGAACTTGATGTAGATCAAA 8973
Db 563 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 622
QY 8974 GGAAGAGTGTGCTGTATGCTGGAATTTGTAATTTGTAATTTGTAATTTGTAATTT 9033
Db 623 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 682
QY 9034 AAGAAAAATGTAGAACCAAAATCGGAAAGCATGCTATTGTAGAACTACTGAAACAA 9093
Db 683 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 742
QY 9094 GAATATCAAGCATTTCAAGAGCAAAAGTAAATATTCTTGAAAGAGAGACGCTGACGT 9153
Db 743 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 802
QY 9154 GCAGTGCATATCGAATGTACACATTTCCATGAGATGATTTAAATTTTGCAAG 9213
Db 803 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 862
QY 9214 CAGTATGCTCTTCTCAATTAATTAACCAAAATTCAAAAATTAATTACTTTAGCATCA 9273
Db 863 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 922
QY 9274 AGTAGGAAATCGAATGTGAATGTTCAATGGGCTGCTGAAGCAAGAGTGCAGAGCAAA 9333
Db 923 AAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 982
QY 9334 GCGACAGTATAGTAAAGATCAAAATTAATGAATTAATTTGATTTAGCAGAGAAA 9393
Db 983 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 1042
QY 9394 ATTAAACAGAGGAAACATCAATGATATGCGGATATGATTAATTAATTAATTAAT 9453
Db 1043 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 1102
QY 9454 AAGACAAATTTCTAAGGCTATTCGAGTGCAGAAAGTCAATGCTGCTGCGCACT 9513
Db 1103 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 1162
QY 9514 GCCACTATTGAAAAATGAAGTAAATTTAATATGCGATCCGAAATTTAAAAATAT 9573
Db 1163 AAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAAAAGAAAA 1222
QY 9574 CTGGCAAGATTGAGAGGAAAGCT-AATTAATAAAGCTCGTATGATCTTAATAGGTAGA 9632
Db 1223 AAGAGAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1282
QY 9633 CTGGTATTAAGGATTAATATATCATGCAATTTCTGAAAAAGCATCAAAAAATTTGACAT 9692
Db 1283 AAGAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAT 1342
QY 9693 TCAATCAAGAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 9723
Db 1343 AAGAACAAACAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1373

```

```

RESULT 5
AG435185/c 1843 bp DNA linear GSS 21-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-313C22.T7, genomic survey
DEFINITION
sequence.
ACCESSION
AG435185
VERSION
AG435185.1 GI:48078248
KEYWORDS
SOURCE
ORGANISM
Mus musculus molossinus (Japanese wild mouse)
Mus musculus molossinus
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T.,
Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K. and
Shiomi, T.
Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
JOURNAL
PUBMED
1574823
REFERENCE
2 (bases 1 to 1843)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Submission
Submitted (17-NOV-2003) Maehira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenho-chou, Tsukumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kunya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyada, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9139
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector : pBAC3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
source
location/Qualifiers
1..1843
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-313C22.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match 1.0%; Score 99.8; DB 10; Length 1843;
Best Local Similarity 31.5%; Pred. No. 3.6e-09;
Matches 469; Conservative 0; Mismatches 1017; Indels 1; Gaps 1;
QY 8237 AAGTTGAGAGAGAAATTTGTCAGAAACAAATCGATTGAATGCAATTTCTTAAAGTGAAG 8296
Db 1825 AANNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1766
QY 8297 GTTGGATGAAGATTAAGTATGCTAAATTTCTCTAGTATCAGAAATGAGAGAGAA 8356
Db 1765 NNAANANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1706
QY 8357 TTGCGGACAGAGAGTACTTCTTACAGCAAAAGTAATCTGAATCCGTAGTTGGT 8416
Db 1705 NNNNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1646
QY 8417 TACGAAGCAAGATTATGAAATTAATGATTACACAAAAAATTAATTTTCAGAAAGTCAATG 8476

```

[illegible][illegible]

```

QY 8614 TCTCACTTCGTCAGAAAGCTTGGCTAAATATGAAATTTATGAAATGTAAGAACT 8673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191 ANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1132
QY 8674 GAGAGAGCTTACTCGAGCGGAAACAGCAGCGCTGAAATTTATCAAAAGTACTACA 8733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1131 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1072
QY 8734 CGAGCATTTGGTTCAGGAAATTGGAAATTGAGATTAATTTAGAAAGATTGCAAGAT 8793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1071 ANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1012
QY 8794 AATACGATTGTAAGTCAACGAGACGAAACCAAGAGAGTCTTGTGCGAAGAAATGT 8853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1011 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 952
QY 8854 ATTTCTGTGAAATTAACAATTTCAGGGGAAACAAATTCATCTGAGATTAAGCCAGA 8913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 951 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 892
QY 8914 ATTGTTGGAACCGAAGTGTAAATGTAGATCTTTGAACTGATGTGATCTACAA 8973
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 891 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 832
QY 8974 GGAAGAGTGTGCTATGTGGAATTGTATTTGAAATGTGATTAATTTGATTT 9033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 831 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 772
QY 9034 AAGAAATATGTAGAGCAAAATCGAGAGCATGTATTTAGAACTACTGAAACAA 9093
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 771 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 712
QY 9094 GAATATCAAGCATTTACAGAGCAAAAGTAATTTCTTGAAAGAGAGCGCTGACCT 9153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 711 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 652
QY 9154 GCAGTGCATATCGAATGTACATTTCCAAATGAGATGATTAATAATTTGGCAAG 9213
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 651 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA--AAAAAAAAAAAAAAAAAAAAAAAAA 595
QY 9214 CAGTATGCACTCTTCAATTAATTAACCAAAATTTCAAAATTAATTTACTTTAGATCA 9273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 594 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 535
QY 9274 AGTAGTCATCGAATGTGAATGTTTCATGGGCTGAGCAAGAGTGCAGAGCCAAA 9333
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 534 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 475
QY 9334 GCGACAGTTAGTGTAAAGATCAATTAATTAAGAACTAATTTGATTTAGCAGAAA 9393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 474 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 415
QY 9394 ATTAATAACAGAGGAAACATCATGTATATCGCGATATGATTAATAATTTAATTAAGT 9453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 414 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 355
QY 9454 AAGACAAATTTAGAGCTATTGGGAGTGCAGAAAGTCAAGCTGCTTGGCAACT 9513
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 295
QY 9514 GGCACATTTGAGAAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 9573
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 235
QY 9574 CTGCAAGATTTGAGAGGAAAGCTAATTAATAAAACGTCGATAGATCTAATCAGTAGAC 9633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 175
QY 9634 TGTATATCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9693
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 115

```

```

QY 9694 CATCAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 85

RESULT 7
AG311072      1615 bp   DNA      linear   GSS 18-DEC-2004
LOCUS
DEFINITION
Mus musculus molossinus DNA, clone:MSMG01-093106.TU, genomic survey
sequence.
ACCESSION
AG311072
VERSION
AG311072.1
KEYWORDS
SOURCE
GSS.
ORGANISM
Mus musculus molossinus (Japanese wild mouse)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Abe, K., Noguchi, H., Tagawa, K., Yuzurika, M., Toyoda, A., Kojima, T.,
Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwa, K. and
Shiroishi, T.
Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
JOURNAL
15574823
PUBMED
2 (bases 1 to 1615)
REFERENCE
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Substitution
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the mouse BAC library MSMG01. For BAC
library availability, please contact Kunyua Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TU
LIBRARY
Vector : PBAC3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
source
1..1615
Location/Qualifiers
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMG01-093106.TU"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMG01 Mouse Male BAC library"
ORIGIN
Query Match 1.0%; Score 97.4; DB 10; Length 1615;
Best Local Similarity 43.1%; Pred. No. 1.1e-08;
Matches 569; Conservative 0; Mismatches 746; Indels 4; Gaps 2;
QY 8409 AGTTGTTTACGAAAGCAAGATTATGAAATATGATTTACACAAATAATATTTTCAG 8468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 AATTGTTTACGAAAGCAAGATTATGAAATATGATTTACACAAATAATATTTTCAG 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8469 AGTCAATGCTCTTGTCTTAAAGATTCAGAAAGATGAAGCAATATGATCTTTAGCGGT 8528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 255
QY 8529 AGCGGTGTCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	256	AAAAAAAAAGAAAAAGAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAAGAAAAAGAAA	315
Qy	8589	TACAACTGTAATGAGAGAAACGTATCTCACTTCGTGCAAAAGCTTTGGCTAAATAATGA	8648
Db	316	GAATAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	375
Qy	8649	AAATATGGAATATGTAAAAAGAACTCGAGAGCCTTAGTCGAGCGGAAACGACGCCCT	8708
Db	376	AAAAAAAAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAGAAAAATTAAGAAGAAAAAAA	435
Qy	8709	TGAAATTAATATCAAAAGACTACTACAGAGCAATTGGTTCAGAGAAATTGGAAATTGAGA	8768
Db	436	TAAAAAATAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAGAAAAATTAAGAAGAAAAAAA	495
Qy	8769	TAAATTAAGAAACGATTGCAAGAGATTAATACAAATTGTAAGATCAACGAGACGAAACCA	8828
Db	496	GAAATTAAGAAAGAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	555
Qy	8829	AGCAGGCTCTGTCCGAAAGAAATGTATTTCTGTGAAAAATCAATTTCAAGGGAGACAAA	8888
Db	556	AATATAAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA	615
Qy	8889	ATCATCCATTGAGATTAAGCCAGAAATTGTTGAAACCGAAGGTAAATGTAGATGCTTT	8948
Db	616	AAAGAAAAA	675
Qy	8949	GAATGAACCTGTATGTATGATCTACAGAGAAAAAGTGGGCTATGGTGAATTGTAATTGG	9008
Db	676	AAAGAAATTAATAAAAAAAAAAAAA	735
Qy	9009	AAATGTGATGTAATTAATGTGATTAGAAAAATGTAG-AAAGCCAAATCGGAAGACATG	9067
Db	736	AAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAATTAATAAGAAAAAAATTAATAAAAAAGAA	795
Qy	9068	CTATGTGAAACTACTGAGAAACAAGAAATCAACGATTTACAGACCAAAAGTAATTA	9127
Db	796	AAAAATTAATTAATAAAAAAAAAAGAAATGAAAAAAATTAATAAAAAAAAAAAAAAAAAAG	855
Qy	9128	TTCTTGGAAAAAGAGACGCTGACGCTGAGCTGCATATCGAATGTACACATTTCCATG	9187
Db	856	AA	915
Qy	9188	AGATGATATTAAAAATTTGGCAAGCATATGATCTTCTCAATTAATTAACCAAAATTT	9247
Db	916	GAATAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAATTAATAAAAAAAAAAAAAAAAAAAA	975
Qy	9248	CAAAATTAATTTACTTAGCATCAAGTATGCAATGTAAGTGAATGTCATGGCGGTG	9307
Db	976	AAAAAAAAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAGAAAAATTAAGAAAAAAATTA	1035
Qy	9308	CTGAAGCAGAGCTCAGAGGCCAAAGCGACAGTTAGTGTAAAGATCAATTAATTAAGAA	9367
Db	1036	ATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAGAAAAAAATTAATAAAAAAAAAAG	1095
Qy	9368	CTAATTAATGTTGATTAGCAGAAAAATTTAAACAGAGGAACATCATGTATATGCGG	9427
Db	1096	AAAAAAAAAGAAAAAAATTAATAAAAAAAAAAAAAAAAAATTAAGAAAAAAATTAATA	1155
Qy	9428	GATATGATTAATAATTAATATATTAATTAAGACAAATTTCAAGGCTATTGCGG---ATGCCA	9484
Db	1156	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAATTAATAAAAAAAAAAGAAAAAAA	1215
Qy	9485	AAAGTCATGCTGCAGCTGCTTGGCAACTGCCATTAATGAAAAAATGAAGTAATTTTA	9544
Db	1216	AAAGAA	1275
Qy	9545	ATTAATGCGATCGAGATTTTAAATTAATCTGCAAGATTTGGAAGGAAAGCTAATAAA	9604
Db	1276	AAAAATTAATAAAAAAAAAAGAAATGAAGAGAAAAAGAAAAAGAAAAAGAAAAAGAA	1335
Qy	9605	AAACGTCCGTAGATCTAATCAGGTAGACTGGTATACGATTAATATATCATGCACTTCTT	9664

[illegible]

OY	8601	GGAGGAAACGATATCTCACTTCGCGCAAAAGCTTGGCTAAATATGAAATTTATGAAA	8660
DB	1195	NAAAAANANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	1136
OY	8661	TGTAAAGAAAGAACTGAGAGACCTTAGTCCGAGCCGAAACAGACCGCTTGAATTTATAC	8720
DB	1135	ANAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA	1076
OY	8721	AAAGAGTACTACAGAGCAATGGTTGAGAAATTTGGAAATTTGAGATAAATTAGAAAC	8780
DB	1075	AAAAAANANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA	1016
OY	8781	GATTGCAGAGATATACTAGTTGTAGACTCAACGAGACGGAACCAAGAGCTCTTGT	8840
DB	1015	AA	956
OY	8841	CGAAGAGATGATTTCTGTGAATAATACATTTCAAGGGAAACAAATCATCATTTGA	8900
DB	955	AA	896
OY	8901	AGATTAAGCCAGAAATTGTTGAAACCGAAGTGAATATAGATGCTTTGAATGAACCTTGA	8960
DB	895	ANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA	836
OY	8961	TGTAGATCTCAAGAAAAAGTGGCTATGTGGCAATTGGTATTTGAAATGTTGATGT	9020
DB	835	AA	776
OY	9021	AAATTAATGTGATTAAGAAAAATGTAGAACCAAAATCGAAGACATCTATTTGTGAAC	9080
DB	775	AA	716
OY	9081	TACTGGAAGAACAGATATACAGACTTTACAAAGCAAAAGTAATTTCTGAAAAAG	9140
DB	715	AA	656
OY	9141	AGAGCTGACGCTGAGCTGCAATATCGAATGTACACTTTTCAATGATGATATTAA	9200
DB	655	AA	596
OY	9201	AAATTTGGCAAGCAGTATGCACTTCTCAATTAAACCAAAATTCAAAAATTAATAT	9260
DB	595	AA	536
OY	9261	TACTTTAGCATCAAGTAGTAGATCGAATGTGAATGTTCAATGGGTGGCTGAAGACAG	9320
DB	535	AA	476
OY	9321	TGCAGAGCCAAAGCAGATTAGTGAAGATCAATAAATAGAACTAATAATGTTGA	9380
DB	475	AA	416
OY	9381	TTTAGCGAGAAAAATTAAACAGAGGGAACATCAATGTATATGCCGATATGTAAAAA	9440
DB	415	AA	356
OY	9441	TTATATATTAAGTAGACAATTTCTAAGGCTATTGCCGATCCAAAAGTCATGCTGCAGC	9500
DB	355	AA	296
OY	9501	TGCTTCGCAACTGCCACTATTGAATAAATGAAGTAATAATTATATATGCGATCCGAGA	9560
DB	295	AA	236
OY	9561	ATTTAAATTAATTCGCGACAGTTGGAAAGGAAAGCTAATAAAAACGTGGTAGATC	9620
DB	235	AA	176
OY	9621	TAAATCAGTAGCTGTATACGGAATTAATATACATGCACTTCTGTGAAAAAGCATCAA	9680
DB	175	AA	116
OY	9681	AAATATGACATATCAATCAAGAGAGAAAAAGGAAAAA 9723	

Db	115	AAA	73
RESULT 9			
AC386893/c			
DEFINITION	AG386893	1519 bp	DNA
LOCUS	Mus musculus molossinus	DNA, clone:MSMg01-201B07.TJ, genomic survey	GSS 21-DEC-2004
ACCESSION	AG386893		
VERSION	AG386893.1	GI:47998098	
KEYWORDS	GSS.		
SOURCE	Mus musculus molossinus	(Japanese wild mouse)	
ORGANISM	Mus musculus molossinus		
REFERENCE	Eularyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murine; Mus.		
AUTHORS	1 Abe, K., Noguchi, H., Tagawa, K., Yuzurika, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwa, K. and Shiroishi, T.		
TITLE	Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC-end		
JOURNAL	Genome Res. 14 (12), 2439-2447 (2004)		
PUBMED	2 (bases 1 to 1519)		
REFERENCE	15574823		
AUTHORS	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Saitoh-chou, Tsukuba, Ibaraki, 305-0858, Japan		
COMMENT	(E-mail: hattori@gscc.riken.jp, URL: http://ngp-gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)		
	Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kunia Abe (abe@cc.riken.jp).		
	Tsukuba Institute, Bio Resource Center,		
	The Institute of Physical and Chemical Research (RIKEN) 3-1-1		
	Koyadai, Tsukuba, 305-0074 Japan		
	phone: 81-298-36-9189, fax: 81-298-36-9199		
	e-mail: abe@cc.riken.jp		
	PRIMERS		
	Sequencing : TJ		
	LIBRARY		
	Vector : pBACe3.6		
	R Site 1 : ECORI		
	R Site 2 : ECORI		
FEATURES	Location/Qualifiers		
SOURCE	1..1519		
	/organism="Mus musculus molossinus"		
	/mol_type="genomic DNA"		
	/sub_species="molossinus"		
	/db_xref="taxon:57486"		
	/clone="MSMg01-201B07.TJ"		
	/sex="male"		
	/tissue_type="mixture of kidney and spleen"		
	/clone_lib="MSMg01 Mouse Male BAC library"		
ORIGIN			
Query Match	1.0%	Score 96.8	DB 10; Length 1519;
Best Local Similarity	38.8%	Pred. No. 1.4e-08;	
Matches 500; Conservative	0;	Mismatches 789; Indels 0; Gaps 0;	
0Y	8435	AAAAATATGATTACACAAAAAATATTTTCAGAGTCATGCTGCTTTAAATGATA	8494
Db	1512	ANNAN	1453
0Y	8495	CAAGAATGAGCAGATATAGATCTTTAGCGGTAGCGGTGTCATGCACAGAGCAACA	8554
Db	1452	NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANANANANANANANANANANANAN	1393
0Y	8555	ACAAAGATTTCAGAGTCAACACAGTTAACTTCTACACTGTAATGAGAAACGTAT	8614

```

Db      1392 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1333
Qy      8615 CTCACCTTGTCGAAAGCTTTGCTAAATAATGAAATTTATGAAATGTAAAGAGACTG 8674
Db      1332 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1273
Qy      8675 GAGAGCCTTAGTCGAGCGGAAACAGACCGTTGAAATATACAAAGACTACTACAG 8734
Db      1272 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1213
Qy      8735 GAGCATTGTTGCGAGAAATTTGGAAATTTGAGATTAATTAGAAACGATTGCAAGATA 8794
Db      1212 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1153
Qy      8795 ATACGATTGTAAGAGTCAACGAGACGAAACGAGAGTCTGTGCGAAAGATGTA 8854
Db      1152 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1093
Qy      8855 TTTCTGTGAATAATACATTTTCAGGGGAAACAAATCATCTGAAGATAAGCCAGAA 8914
Db      1092 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1033
Qy      8915 TTGTTGAAACCGAGAGTGTAAATGTGATGCTTTGATGATGCTTGATGTAGATCTACAG 8974
Db      1032 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 973
Qy      8975 GAAAAAGTGCGCTAGTGTGAAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTT 9034
Db      972 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 913
Qy      9035 AGAAAAATGTAGAGCCAAATTCGAGACATGCTATTGTAGAACTACTGAGAAACAAG 9094
Db      912 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 853
Qy      9095 AATATCAGACATTTACAAGACCAAGTAAATTTCTTGGAAAGAGACGCTCCAGCTG 9154
Db      852 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 793
Qy      9155 CAGCTGCAATATGCAATGTACATTTCCATGAGATGATTAATAATTTGGCAAGC 9214
Db      792 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 733
Qy      9215 AGTATGCACTTCTCAATTAATTAACCAAAATTTCAAAATTAATTAATTTAGCATCA 9274
Db      732 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 673
Qy      9275 GTAGTAAATCGAATGTGATTTTCATGGGGTGGCTGAGACAGAGTGCAGAGCCAAAG 9334
Db      672 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 613
Qy      9335 CGACAGTTAGTAAAGATCAATAATAGAACTAATATGTTGATTTAGCAGAGAAA 9394
Db      612 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 553
Qy      9395 TTAATAACAGAGGAAACATCAATGTATATGCCGAGATATGATAAATAATTAATTAAGTA 9454
Db      552 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 493
Qy      9455 AGACAAATTTCTAAGGCTATTGGCGATGCCAAAGTCACTGCTGCTGGCACTG 9514
Db      492 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 433
Qy      9515 CCACATTTGAAATAATGAAATTTTAATAATGCAATCGAGATTTAAATAATTAATTC 9574
Db      432 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 373
Qy      9575 TGCAAGATTGAGAGGAAAGCTAATTAATAAAGCTCGTAGGATCTAATCAGTAGACT 9634
Db      372 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 313
Qy      9635 GGTATACGATTAATATATACATGCGATTTCTTGAAGAAAGCATCAAAATTTGACATATC 9694
Db      312 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 253

```

```

Qy      9695 AATCAAGAGAGAGAAAAAGGAAAAA 9723
Db      252 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 224

RESULT 10
DR124185
LOCUS
DEFINITION
pseudobscura cDNA clone M7 5', mRNA sequence.
ACCESSION
DR124185
VERSION
DR124185.1 GI:67840883
KEYWORDS
EST.
SOURCE
Drosophila pseudoobscura
ORGANISM
Drosophila pseudoobscura
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1691)
Richards,S., Liu,Y., Bertencourt,B.R., Hradecky,P., Letovsky,S.,
Nielsen,R., Thornton,K., Hubisz,M.J., Chen,R., Meisel,R.P.,
Coutome,O., Hua,S., Smith,M.A., Zhang,P., Liu,J., Bussemaker,H.J.,
van Batenburg,M.F., Howells,S.L., Scherer,S.E., Soderstrom,E.,
Matthews,B.B., Crosby,M.A., Schroeder,A.J., Ortiz-Barrientos,D.,
Rives,C.M., Metzker,M.L., Muzny,D.M., Scott,G., Steffen,D.,
Wheeler,D.A., Worley,K.C., Havlak,P., Durbin,K.J., Egan,A.,
Gill,R., Hume,J., Morgan,M.B., Miner,G., Hamilton,C., Huang,Y.,
Walton,L., Verdusco,D., Clerc-Blanchard,K.P., Dubchak,I.,
Noor,M.A., Anderson,M., White,K.P., Clark,A.G., Schaeffer,S.W.,
Gelbart,W., Weinstein,G.M. and Gibbs,R.A.
Comparative genome sequencing of Drosophila pseudoobscura:
chromosomal, gene, and cis-element evolution
Genome Res. 15 (1), 1-18 (2005)
15632085

JOURNAL
PUBMED
COMMENT
Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephen@bcm.tmc.edu
NCBI Trace Archive: 226706818
Insert length: 1750 Std Error: 0.25.
Location/Qualifiers
1..1691
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="M7"
/dev_stage="0-18h embryo"
/clone_lib="Drosophila pseudoobscura embryonic cDNA
library"
/note="Vector: pOTB7, Site_1: EcoRI; Site_2: XhoI, oligo
dt priming from poly A+ RNA, directionally cloned"

ORIGIN
Query Match 1.0%; Score 96.6; DB 8; Length 1691;
Best Local Similarity 42.1%; Pred. No. 1.6e-08;
Matches 549; Conservative 0; Mismatches 754; Indels 0; Gaps 0;

Qy      8421 AAGCAGATTATGAAATTAATGATTACACAAAAAATATATTTCAGAGTCAATGCTCT 8480
Db      364 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 423
Qy      8481 TGCTTTAAATGATCAAGATGAAGCGAATTTAGAAATTTTAAAGCGGTGCGA 8540
Db      424 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 483
Qy      8541 TGCAAGAGAAACAACAAAGATTACGAGTCAACAAGATTACCTTACAACTGTAA 8600
Db      484 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 543
Qy      8601 TGAGAGAAACGTATCTCAACTTCGTGCAAAAGCTTTGGCTAAATAATTAAGAAA 8660

```

```

Db      544 AGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAATAGAAAA 603
      8661 TGTAAAGAACTGAGAGCCCTTATGTCGAGCGGAAACGCCCGCTGAAAAATTATAC 8720
      604 AAAAAAAAAAAAAAGAAAAAGAAAAAGAAATTAATAAAGAAAAAGAAAAAGAAAA 663
      8721 AAGAGTACTACGAGCATTTGGTTCAGAGAAATTGGGAAATTGAGATAATTAGAAC 8780
      664 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 723
      8781 GATTGCAAGATATACGATTGTAAAGATCAACGAGACGAAACCAAGAGCTTGT 8840
      724 AAAAAAGAAAAAGAAAAAGAAATTAAGAGAGAGAGAGAGAAAAAGAAAAAGAAAAAG 783
      8841 CGAAAAAGATGTATTTCTGTGAAAAATCAATTTCAAGGAAAAACAATCATCATGCA 8900
      784 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 843
      8901 AGATTAAGCCGAATTGTTGAAACCGGAGCTGTAATGTAGATGCTTGAATGAACTGA 8960
      844 AAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 903
      8961 TGTAGATCTACAGAAAAAGTGTGTGCTATGCTGAAATTTGTAATGTTGATGT 9020
      904 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 963
      9021 AATTAATGTGATTAGAAAAATAGAACCAAAATCGAAGACATGCTATTGTAGAAC 9080
      964 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1023
      9081 TACTGAAAAACAATATATCAACATTTACAGAGCAAAAGTAAATTTCTTGAAAAAG 9140
      1024 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1083
      9141 AGAGCTGCACTGAGCTCAATATCGAATGTACACATTTCCATAGAGATGATTTAA 9200
      1084 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1143
      9201 AATTTGGCAAGCACTATGCTCTCTCAATTAATACCAAAATTCAAAAATATATAT 9260
      1144 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1203
      9261 TACTTACGATCAAGTAGTAATCGAATGTGAAATGTCAGGGGTGCTGACAGAGG 9320
      1204 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1263
      9321 TGCAGAGCCAAAGCAAGTATGTAAGAAATCAAAATTAATAGAACTAATATGTTGA 9380
      1264 AGAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1323
      9381 TTTAGAGAAAAATTAAGAACAGAGGAAACATCAATGTATATCCGAGATATATAAAA 9440
      1324 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1383
      9441 TTTAATATTAAGTAACAAATTTCTAAGGCTATTTGCGAGATGCCAAAGTCAATGCTGACG 9500
      1384 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1443
      9501 TGCTTCGCAACTGCCCTATTGAAAAAATGAAGTAAATTTAATTAATCGATCCGAGA 9560
      1444 AATTAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1503
      9561 ATTAAAAATTAATCTGCGAAGATTGGAAGGAAAGCTAATTAAGAAACGTGCTAGAGTC 9620
      1504 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1563
      9621 TAATCAGTGAATCTGGATACGATTAATATATCATGCAATTTCTTGAAAAAGCATACAA 9680
      1564 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1623
      9681 AAAAAATGACATATCAATCAAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 9723

```

```

Db      1624 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1666

RESULT 11
LOCUS   CL118787
DEFINITION
  CL118787 2001 bp DNA linear GSS 05-JAN-2004
  ISB1-72M23 T7.2 ISB1 Xenopus tropicalis genomic clone ISB1-72M23,
  genomic survey sequence.
ACCESSION
  CL118787
VERSION
  CL118787.1 GI:40612422
KEYWORDS
  SOURCE
  ORGANISM
    Xenopus tropicalis (western clawed frog)
    Eukaryota; Metazoa; Chordata; Vertebrata; Eumelostomi;
    Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
    Xenopodinae; Xenopus; Silurana.
REFERENCE
  1 (bases 1 to 2001)
  Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
  Mardis, E. and Wilson, R.
  A physical map of the xenopus tropicalis genome
  Unpublished (2003)
  CONTACT
    JOURNAL
    COMMENT
    Genome Sequencing Center
    Washington University School of Medicine
    Email: submissions@watson.wustl.edu
    Insert Length: 75000 Std Error: 0.00
    Seq primer: T7 TAATACGACTCACTATAGG
    Class: BAC ends
    High quality sequence start: 384
    High quality sequence stop: 783.
    Location/Qualifiers
      source
        1..2001
        /organism="Xenopus tropicalis"
        /mol_type="genomic DNA"
        /db_xref="taxon:8364"
        /clone="ISB1-72M23"
        /clone_1lb="ISB1"
        /note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
        Library Segment 1"

ORIGIN
Query Match 1.0%; Score 96.2; DB 10; Length 2001;
Best Local Similarity 42.8%; Pred. No. 2e-08;
Matches 558; Conservative 0; Mismatches 736; Indels 11; Gaps 2;

      QY      8422 AAGCAAGATTGAAAAATATGATTAACACAAAAAATATATTTCAGAGTCAATGCTCTT 8481
      DB      181 AAAAAAAAAAAAAAAAAANNNNNNNNAAAAAGAAAAACCTTATTAAGGGGNNNNNNNA 240
      QY      8482 GCTTTAAATGATACAAAGATGAAGCAATATGAAATCTTTAGCGGTAGCCGTGCGAT 8541
      DB      241 NNAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 300
      QY      8542 GCACAGAGAAACAAGACATTTAGAGATCAACAGTTAACTTCAACTGTAAAT 8601
      DB      301 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 360
      QY      8602 GGAAGAAAGTATCTCACTTCGTGCAAAAGCTTTGGCTAAAAATGAAAAATTATGAAAT 8661
      DB      361 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 420
      QY      8662 GTTAAAGAACTGAGAGCCCTTATGTCGAGCGGAAACGACCGCTGAAATTTTACA 8721
      DB      421 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 480
      QY      8722 AAGAGTACTACGAGCATTTGGTTCAGAGAAATTGGGAAATTGAGATAATTAGAAACG 8781
      DB      481 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 540
      QY      8782 ATTGCAAGATTAATCGATTGTAGAGTCAACGAGACGAAACCAAGAGAGCTTTGTC 8841
      DB      541 AAAAAATTAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAATTAAG 600

```

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Xenopodinae; Xenopus; Silurana. 1 (bases 1 to 1547)	Kremnitzki, C., Carter, J., Mcherson, J., Warren, W., Graves, T.,	Wardis, E. and Wilson, R.	A physical map of the xenopus tropicalis genome	Unpublished (2003)
	Contact: Richard K Wilson			

Washington University School of Medicine
Email: subunits@wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq Primer: Sp6 ATTACGTACACTTAG
Class: BAC ends
High quality sequence start: 909
High quality sequence stop: 1073.

```

FEATURES
  source
    location/Qualifiers
      1.1547
        /organism="Xenopus tropicalis"
        /mol_type="genomic DNA"
        /strain="Nigerian frog"
        /db_xref="taxon:8364"
        /clone="CH216-108P4"
        /sex="male"
        /cell_line="Stock 248 F7A2, inbred N7"
        /clone_lib="CH216"
        /note="Vector: pPARAC2.1; CHO1-216 Xenopus tropicalis
BAC library"

```

ORIGIN

Query Match	1.0%	Score 95.4	DB 10	Length 1547
Best Local Similarity	42.5%	Pred. No. 2.8e-08		
Matches 555; Conservative	0	Mismatches 749	Indels 2	Gaps 1

Qy	8421	AAAGCAAGATTATGAAAAATAATGATTACACAAAAAATATATTTGAGAGTCAATGCTCT	848
Db	173	AAAAAAAAAAAAAAAAAGAAAAACCCCATTAATTAAGAAAAACNNAAAAAAAAAAAAAG	232
Qy	8481	TGCTTTAATGATCAAGAAATGAAAGCAATTAATAATCTTACGGGTAGCCGGTGTGCA	8541
Db	233	AAAAAAAAAAAAAAAAAGAAA	292
Qy	8541	TGCACAAAGAACAAACAAAGCATTTACAGATCAACAAAGTTAACTTCAACACTGTAA	8601
Db	223	AAAGAAAGAAA	352
Qy	8601	TGGAGAAACGTATCTCACTTCGTCAAAAGCTTGCTTAATAATGAAATTATGAAA	8661
Db	353	AAGAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAA--AAAAAA	410
Qy	8661	TGTAAAGAACTGGAGAGCCTTAGTGGAGCGGAAACAGCAGCCGTTGAATAATTATAC	8721
Db	411	AAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAA	470
Qy	8721	AAAGGTACTACAGAGCATTTGGTTGCAGAAATTGGAAATTGGAGATTAATTGAAAC	8781
Db	471	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAGAGAAAAAAAAAAAAAAAAAAAA	530
Qy	8781	GATTGCAGAGATTAATGACATTGTTAAGACTCAACGAGACGGAACCAAGAGAGCTTGT	8841
Db	531	AAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAA	590
Qy	8841	CGGAAGAGATGATTTCTGTGAAATAATACATTTCAAGGAGAAACAAATCATCATTTGA	8901
Db	591	AAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAGAGAGAAAAAAAA	650
Qy	8901	AGATTAAGCCGAATTTGTGGAACCGAGAGTGTAAATGTAGATGCTTTGAATGAACCTTA	8961
Db	651	AAAAAAAAAAAAAAAAAAAAAAAACAAAAGAGAAAAAAAAAGAGAGAAAGAAAAAAAA	710
Qy	8961	TGTAGATCTACAGAGAAAGAGGTGGCTATGTGTGAATTTGGTATTTGGAAATGTTGATGT	9021
Db	711	AGAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAGAGAGCAAAAAAAAAAAAAAAAAAGAAA	770

```

QY 9021 AATATAATGATTAGAAAAATGTAGAGCCAAATCGAAGACATGCTATTGTAGAAC 9080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 771 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 830
QY 9081 TACTGAAACAAGAATATCAGCATTTACAGAGCAAAAGTAATATTTCTGAAAAGG 9140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 831 AAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 890
QY 9141 AGACGCTGCACTGCTGCTGCAATTCGAATGTAACATTTCCAAATGAGATGATATTAA 9200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 891 AAAAAAAAAAGAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAA 950
QY 9201 AATTTGGCAAGACAGTATGATCTTCTCAATTAATTAACCAAAATTTCAAAATATATAT 9260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 951 AGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1010
QY 9261 TACTTTAGCATCAAGTAGAATCGAATGTAATGTTATGAGGCTGAGCAAGAG 9320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1011 AAAAAAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1070
QY 9321 TCGAGAGCCAAAGCAGCACTTATGTTAAAGAAATCAATTAATTAAGAACTTAATATGTA 9380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1071 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1130
QY 9381 TTATGACAGAAAAATTAATAACAGAGGAAACATCAATGTAATGCGGATATGATATAA 9440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1131 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1190
QY 9441 TTATTAATTAAGTAGAACAATTTCTAAGCTATTGCGATGCCAAAGTCATGCTGACG 9500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1250
QY 9501 TCGTTCGCACTGCCATCTTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1251 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1310
QY 9561 ATTTAATAATTAATCTGCAAGATTTGAAAGGAAAGCTAATTAATAAAGCTCGTAGATC 9620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1311 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1370
QY 9621 TAATCAGTAGACTGTTATTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1371 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1430
QY 9681 AAAAAATGACATTCATCAAGAGAGAGAAAAAGGAAAAATGA 9726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1431 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1476

```

```

RESULT 13
LOCUS CL081992 1784 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-165M9_RM4.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CL081992
VERSION CL081992.1 GI:40537905
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
            Xenopodinae; Xenopus; Silurana.
            1 (bases 1 to 1784)
            Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
            A physical map of the xenopus tropicalis genome
            Unpublished (2003)
            Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submls@wustl.edu
            Insert Length: 175000 Std Error: 0.00
            Seq primer: RM4 ctcaaggcgatcgctgcgagc

```

Class: BAC ends
High quality sequence start: 409
High quality sequence stop: 503.
Location/Qualifiers

FEATURES

source

1.1784

/organism="Xenopus tropicalis"

/mol_type="genomic DNA"

/strain="Nigerian frog"

/db_xref="taxon:8364"

/clone="CH216-165M9"

/sex="male"

/cell_line="Stock 248 F7A2, inbred N7"

/clone_1ib="CH216"

/note="Vector: pTARBAC2.1; CHOR1-216 Xenopus tropicalis

ORIGIN

Query Match 1.0%; Score 95.2; DB 10; Length 1784;
Best Local Similarity 43.6%; Pred. No. 3.1e-08;
Matches 515; Conservative 0; Mismatches 664; Indels 2; Gaps 2;

```

QY 8544 ACAGAGACAAACAAAGCAATTAACGATCAACCAATTAATTAATTAATTAATTAATTAATTA 8603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 361
QY 8604 AGCAACGATCTCACTTCGTCGCAAAAGCTTTGCTAAATAATGAATAATTAATGAATATG 8663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 AGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 421
QY 8664 AAAAAAGCACTGAGAGAGCCTTAGTGTGCGAGCGAAACAGACGCCGTTGAATAATTAATCAAA 8723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 481
QY 8724 GAGTACTACAGAGCATTTGTTGCAAGAAATTTGGAAATTTGAGATTAATTAAGAACAT 8783
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 541
QY 8784 TCGAAGAGTAATTAACATTTGTAAGATCAACGAGACGAGAACCAAGAGAGCTTTGTCGG 8843
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 601
QY 8844 AAAAAAGTATTTCTGTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8902
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 661
QY 8903 ATTAAGCCAGAAATTTGTTGAAACCGAAGTGAATGTAGATGCTTTGAATGAATTGATG 8962
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 721
QY 8963 TAGATCTACAGAGAAAAAGTGTGCTATGTCGAAATTTGGTATTTGGAATTTGATGTAA 9022
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 722 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 781
QY 9023 ATTAATGTAATTAAGAAAAATTTGAAGCAAAATGGAAGACATGCTAATTTGTGAAGAACTA 9082
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 782 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 841
QY 9083 CTGAGAAACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 842 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 901
QY 9143 AGCTGACGCTGACGCTGCAATTAATGTAATGTAATTTCAATGAGATGATTAATAA 9202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 961
QY 9203 ATTTGGCAAGACAGTATGATCTTCTCAATTAATTAACCAAAATTTCAAAATTAATTAATTA 9262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 962 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1021
QY 9263 CTTTAGCATCAAGTAGAATGCAATGTAATGTTCAATGCGGTGCTGAAGCAAGAGCTG 9322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1022 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1081

```


[illegible]

RESULT 15
AG332292/c

LOCUS	AG3332292	1482 bp	DNA	linear	GSS 18-DEC-2004
DEFINITION	Mus musculus molossinus DNA, clone:MSMG01-122G19.TJ, genomic survey				

sequence.	
ACCESSION	AG332292
VERSION	1
ORIGIN	AT 470000000

```

VERSION      AG334292.1  GI:47905602
KEYWORDS     GSS.
COMMENT      Mus musculus melanogaster (Transcript 114)

```

SOURCE
mus musculus molossinus (japanese wild mouse)
ORGANISM
mus musculus molossinus
Eukaryota; Metazoa; Chordata; Mammalia; Euteleostomi;
Eulipotyrychia; Rodentia; Muridae; Murinae; Murini;

[illegible]

REFERENCE

1. Abe, K., Nomuchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Koijima, T., Otsuogouchi, T., Matsuo, N., Matsuo, N.

Erawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriaki, K. and Shirolshi, T.

TITLE Contribution of Asian mouse subspecies *Mus musculus molossinus* to genomic constitution of strain C57BL/6J, as defined by BAC-end

Sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)

PUBMED 15574823
REFERENCE 2 (bases 1 to 1482)

AUTHORS	TITLE
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.	Direct Submission

Submitted (17-NOV-2003) Masahito Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

I-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: <http://hgp.gsc.riken.go.jp/>)

COMMENT
clones are derived from the mouse BAC library MSMG01. For BAC
information visit <http://www.mcgill.ca/medgen>

For further information and library availability, please contact kunhiya Abe (abe@crc.riken.jp), Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1

the Institute of Physical and Chemical Research (RIKEN), 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-35-6188 fax: 81-298-35-6188

Phone: 02-236-36-2201, Fax: 02-236-36-2155
e-mail: abe@rtc.riken.jp
PRIMERS

Sequencing : TV
LIBRARY

Vector : pBACe3.6
R.Site 1 : EcoRI

R.Site 2 : ECORI.

FEATURES		Location/Qualifiers
SOURCE		1..1482
		/organism="Mus musculus mojosinus"
		/mol_type="genomic DNA"
		/sub_species="mojosinus"
		/db_xref="taxon:57486"
		/clone="MSMG01-120G19.TJ"
		/sex="male"
		/issue_type="mixture of kidney and spleen"
		/clone_lib="MSMG01 Mouse Male BAC library"
	ORIGIN	
	Query Match	1.0%; Score 93.9; DB 10; Length 1482;
	Base Local Similarity	40.1%; Pred. No. 5.8e-06;
	Matches 434; Conservative	0; Mismatches 649; Indels 0; Gaps 0;
Qy	8641	AAAAATGAAATTTGTGGAATGTAAAGAACTGAGAGCGCTTATGTCGAGCGAATAA 8700
Db	1387	AA 1328
Qy	8701	GCAGCCGCTTGAAATTTATCAAGAGTACTACAGAGCATTTGTTGCAGAAATTGGAA 8760
Db	1327	AA 1268
Qy	8761	ATTGGAATTAATTAGAAACGATTCGACAGATTAATCGATTGTAAAGTCAACGAGAC 8820
Db	1267	AAN 1208
Qy	8821	GAACCAAGAGGCTCTTGCGAAGATGCTATTCTGTGAATAATTAATTTCAGCG 8880
Db	1207	AA 1148
Qy	8881	GAACCAAAATCATCATTTGAAAGTAAAGCCAGATTGTTGAAACCGAAGTAAATGTA 8940
Db	1147	NNAA 1088
Qy	8941	GATCCTTGAATGACTTGTATGATCTACAGAGAAAAGTGTGCTATGTTGAATT 9000
Db	1087	AA 1028
Qy	9001	GCTATTGGAATGTTGATGTAAATATGTGATTTAAGAAAATGTAGAACCAAAATCGA 9060
Db	1027	AA 968
Qy	9061	AGACATCTATTGTAGAACTACTGGAACCAAGATATCAAGCATTTACAGAGCAAAA 9120
Db	967	AA 908
Qy	9121	GTAATATTTCTGGAAAGAGACGCTGACGCTGACATGTAATGTAACATT 9180
Db	907	AA 848
Qy	9181	TCCATGATGATGATTAATAATTGGCAAGCAGTATGCTTTCTCAATTATATACC 9240
Db	847	AA 788
Qy	9241	AAAAATTCAAAATAATATATTACTTTGACATCAAGTAGTAATGGAATGTTTCAT 9300
Db	787	AA 728
Qy	9301	GGGCTGCTGACAGCAAGAGTGCAGAGCCAAAGCGACAGTTAGTGAAGATCAATA 9360
Db	727	AAN 668
Qy	9361	AATGAACTAATAATGTTGATTAGCAGAAAATTTAAACAGAGGAAACATCAATGTA 9420
Db	667	AAANNA 608
Qy	9421	TATGCGGATATGATAAATAATTAAATATAGTAGACAATTTAAGGCTATTCGGAT 9480
Db	607	AA 548
Qy	9481	GCCAAAAGTCATGCTGACGCTGCTTCGGCACTGCCATTTGAAAAAATGAATAAA 9540

[illegible][illegible]

Db	1193	AAATANNAAAAAANANANNNNTAAAAAAGNNNNAAAAAANAAAAA	1252
Qy	9603	AAAAACGTCGGTAGCATTAATCAGCTGACCTGTATCGGATTAATATACATGCGATTC	9662
Db	1253	NAAAAAANANANANNAATAAAAAANATTAATAAAAAGAAAAAANAAAAAATA	1312
Qy	9663	TTCTGAAGAACCATCAAAATAATTGACATATCAATCAAGAGAGAGAAAA	9714
Db	1313	NTAANAAAAAANAAAAAANNAATNAAAAAANAAAAAANAAAAAAGAAA	1364

RESULT 17	CG748014	1531 bp	DNA	linear	GSS 24-OCT-2003
LOCUS	CG748014				
DEFINITION	P041-4-B02.za Ppa EcORI BAC Library Pristionchus pacificus genomic.				
ACCESSION	CG748014				
VERSION	CG748014.1				
KEYWORDS	GSS.				
SOURCE	Pristionchus pacificus				
ORGANISM	Pristionchus pacificus				
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;				
AUTHORS	Mediologasteridae; Pristionchus.				
TITLE	1 (bases 1 to 1531)				
JOURNAL	Srinivasan,J., Sinz,W., Jeesse,T., Wiggers-Perebolte,L., Jansen,K.,				
PUBMED	Buntjer,J., van der Meulen,M. and Sommer,R.J.				
COMMENT	An integrated physical and genetic map of the nematode Pristionchus pacificus				
	Mol. Genet. Genomics 269 (5), 715-722 (2003)				
	12884007				
	Contact: Sommer RJ				
	Evolutionary Biology				
	Max-Planck-Institute for Developmental Biology				
	Spemannstr. 37-39, Tuebingen D-72076, Germany				
	Tel.: 00497071601371				
	Fax: 00497071601498				
	Email: ralf.sommer@uebingen.mpg.de				
	Class: BAC ends.				
FEATURES	location/Qualifiers				
source	1..1531				
	/organism="Pristionchus pacificus"				
	/mol_type="genomic DNA"				
	/strain="California"				
	/db_xref="taxon:54126"				
	/clone_lib="Ppa EcORI BAC library"				
	/note="The library was generated by a partial digest of the genomic DNA with EcORI and cloning into the BAC vector."				
ORIGIN					
Query Match	1.0%; Score 93.2; DB 10; Length 1531;				
Best Local Similarity	37.7%; Pred. No. 7,7e-08;				
Matches	491; Conservative 0; Mismatches 812; Indels 0; Gaps 0;				
Qy	8421	AAAGCAAGATTATGAAAATAATGATTACACAAAAAATATATTTTCAGAGTCAATGCTCT	8480		
Db	180	AAANANANANANANNNAAAAAANANANANAAAAAANANANANANAAAAA	239		
Qy	8481	TGCTTAATATATCAAGAAATGAAGCAATATAGATCTTACGGCGTACCGGCTGCA	8540		
Db	240	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA	299		
Qy	8541	TGCAACAAGAACAAACAAGCATTTACAGATCAACAAGTTAACTTCAACAATGTAA	8600		
Db	300	NAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA	359		
Qy	8601	TGAGAGAAACGTATCTCACTTCGTGCAAAAGCTTGGCTAAATAATATATGAAA	8660		
Db	360	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA	419		
Qy	8661	TGTAATAAGAACTGAGAGACCTTAGTGTGGAGCGGAAACAGACGCCGTGAAATATATAC	8720		
Db	420	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA	479		

[illegible]

DEFINITION ISB1-53P23_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-53P23,
genomic survey sequence.
ACCESSION CL110653
VERSION CL110653.1 GI:40604288
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1594)
Kremetzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: subml@wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: Sp6 ATTTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 390
High quality sequence stop: 470.
Location/Qualifiers
1. 1594
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-53P23"
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN

Query Match 1.0%; Score 93; DB 10; Length 1594;
Best Local Similarity 43.1%; Pred. No. 8.6e-08;
Matches 556; Conservative 0; Mismatches 725; Indels 8; Gaps 2;

8417 TACGAGAGAGATTGTAATGATGATTCACAAATTAATATTGAGAAAGTCAATG 8476
1562 TAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1503
8477 CTCTGCTTAAATGATACAAAGATGAGCGAATAGATCTTACGCGTAGCGGTG 8536
1502 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1443
8537 TGCATGCAAGAGCAAAACAAGATTACGAGATCAACAAGTTACTTCTACAACTG 8596
1442 TTAATTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1383
8597 TAAATGAGAGAAACGATCTCACTTCGTCAGAAAGCTTGGCTAATAATGAAATTATG 8656
1382 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1330
8657 GAAATGTAAAGAACTGAGAGAGCTTAGTCGAGCGGAAACAGACGCGTTGAAATT 8716
1329 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1270
8717 ATCAAAAGAGTCTACAGAGATTGGTTCAGAAATTGGAAATTGAGATAATTAG 8776
1269 AATATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1210
8777 AAACGATTGCAAGAGATTAATGATTTGTAAGACTCAACGAGACGAAACGAGAGTC 8836
1209 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1150
8837 TTGTCGGAAGAATGATTTCTGTAATAATCAATTTCAAGCGGAAACAAATCAATCA 8896
1149 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1090
8897 TTGAAGATTAAGCCAGAAATTGTTGGAACGGAAGTAAATGATGCTTTGAATGAAC 8956

Db 1089 AA 1030
Qy 8957 TTGATGTAGATCTACAGAGAAAGTGGTCTATGTGTGATTTGAAATGTGG 9016
Db 1029 AA 970
Qy 9017 ATGTAAATTAATGTGATTAAGAAAAATGTGAAGCCAAATCGAAGACATGCTATTGG 9076
Db 969 AATTAAAAAACAACAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 910
Qy 9077 AAACATCTGGAAGAACAGAAATTCAGACTTTTCAAGACCAAAAGTAATTTCTTGA 9136
Db 909 AAAAAAATTAACAA-CAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 851
Qy 9137 AAGAGAGCGCTGACCTGACGTGCAATTAATGATGACATTTCCATGAGATGATA 9196
Db 850 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 791
Qy 9197 TTAATAATTTGCAAGAGATGATGATCTTCAATTAATAACCAAAATTCAAAAATA 9256
Db 790 ATTAATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 731
Qy 9257 ATATTACTTTAGCATCAAGTATGATGATGATGATGATGATGATGATGATGATG 9316
Db 730 AA 671
Qy 9317 GAGTGCAAGAGCCAAAGCAGTATGATGATGATGATGATGATGATGATGATGATG 9376
Db 670 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 611
Qy 9377 TTGATTTAGCAGAGAAATTAATAAGAGGAAACATCAATGATATGCGGATATGATA 9436
Db 610 ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 551
Qy 9437 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9496
Db 550 AA 491
Qy 9497 CAGCTGCTTCGCACTGCCACTATTTGAATAAATGAAGTAATTAATTAATGATGC 9556
Db 490 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 431
Qy 9557 GAGATTTAAATTAATCTGCAAGATGGAAGGAAAGCTAATTAATAAATCGTGTAG 9616
Db 430 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 371
Qy 9617 GATCTAATCAGTACGATCTGTATGAGATAATTAATTAATTAATTAATTAATTA 9676
Db 370 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 311
Qy 9677 ACAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9705
Db 310 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 282

RESULT 19
CL118709
LOCUS CL118709 1654 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-72115.T7.2 ISB1 Xenopus tropicalis genomic clone ISB1-72115,
genomic survey sequence.
ACCESSION CL118709
VERSION CL118709.1 GI:40612344
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1654)
Kremetzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)

ORIGIN

/clone_lib="ISB1"
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"

Query Match 0.9%; Score 91.8; DB 10; Length 1256;
Best Local Similarity 37.4%; Pred. No. 1.4e-07;
Matches 441; Conservative 0; Mismatches 739; Indels 0; Gaps 0;

8544 ACAGAGACAAACAAAGCATTTTCAGATCAACAGTTAACTTCAACCTGTAATG 8603
2 AA 61
8604 AGGAAACGTATCTCACTTCGTGCAAAAGCTTGGCTAAATAATGAAATG 8663
62 AA 121
8664 AAAAGACCTGACGAGCCTTACTGACGCGAAGACGACCCGTTGAAATTAACAA 8723
122 AA 181
8724 GAGTACTACGAGCATTTGGTTGCGAAATTGGAAATTTGAGATTAATTGAACGAT 8783
182 AA 241
8784 TGCAGAGATTAATACGATTGTAGATCAACGAGACGAAACCAAGAGCTCTTGC 8843
242 AA 301
8844 AAAAGATGTTATTTCTGTGAAAAATCAATTTCAAGGGAACAAATCACTTGAAGA 8903
302 AA 361
8904 TAAAGCCAGATTGTTGAAACGGAAGTGAATGTAGATGCTTGAATGAATGATG 8963
362 AA 421
8964 AGATCTACAGAAAAAGTGTGCTATGATGGAATTTGATTTGAATTTGATGATAA 9023
422 AA 481
9024 TAAATGATTAAGAAAAATGTAGAGCCAAATCGAGACATCTATTGTAGAACTAC 9083
482 AA 541
9084 TGGAAACAGAAATATCAAGCATTTTCAAGAGCAAAAGTAATCTTGGAAAGAGA 9143
542 AA 601
9144 CGCTGACGCTGACGTCATATCGAATGTACACATTTTCAATGAGATGATTAATAA 9203
602 AA 661
9204 TTTGGCAAGACGATGATCTTCTCAATTAATCAACCAAAATTCAAAATTAATATAC 9263
662 AA 721
9264 TTTAGCATCAAGTGAATGCAATGTGAATGTTCAATGGGCTGCTGAAGCAGAGTGC 9323
722 AA 781
9324 AGAGACCAAGCGACGATTAAGTGAAGATCAATTAATTAATTAATGATTT 9383
782 AA 841
9384 AGCAGAAAAATTTAAACAGAGGAAACATCAATGTATATGCGGATATGATTAATAA 9443
842 AA 901
9444 TAAATTAAGTAAACAAATTTCTAAGCTATTGCGGATGCGAAAGTCATGCTGACGTCG 9503
902 AA 961

9504 TTCGCAACTGCGACTATTGAAAAAATGAACTAAATTTAATTAATGCGATCCGAGATT 9563
962 AA 1021
9564 TAAATAATTAATCTGCGAGATTGGAAGGAAAGCTAATTAATAAAGTCGTGATCTAA 9623
1022 AA 1081
9624 TCAGGTAGCTGCTGTACGATTAATTAATCATGATGCTTCTCGAAAAAGCATCAAAA 9683
1082 AA 1141
9684 ATTGACATATCAATCAAGAGAGAGAAAAAGGAAAAA 9723
1142 AA 1181

RESULT 21
LOCUS CL076558 1485 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-139010 Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CL076558
VERSION CL076558.1 GI:40532471
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryote; Metazoa; Chordata; Vertebrata; Suteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;

REFERENCE
AUTHORS
Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
1 (bases 1 to 1485)
Xenopus tropicalis; Silurana.

TITLE
JOURNAL
COMMENT
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Clase: BAC ends
High quality sequence start: 572
High quality sequence stop: 635.
Location/Qualifiers

FEATURES
source 1. 1485
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-139010"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTAKBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN

Query Match 0.9%; Score 91.8; DB 10; Length 1485;
Best Local Similarity 42.8%; Pred. No. 1.5e-07;
Matches 561; Conservative 0; Mismatches 745; Indels 4; Gaps 2;

8422 AAAGCAAGATTATGAATAATGATTAACAAAAAATTTATTTCAAGATCAATGCTCT 8480
111 AA 170
8481 TCGTTAAATGATCAAGAAATGAAGCAATTAATGATCTTTAGCGGTAGCCGTGTGCA 8540
171 AGAAAAAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 230
8541 TGCACAGGACAAACAAAGCATTTTACGAGATCAACAGTTAACTTCTTACAACTGTAA 8600
231 GATAAAAAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 290

```

Qy 8601 TGGAGAAACGTATCTCACTTCGTCGAAAGCTTTGGCTAAATAATTATGAAA 8660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 291 AAGGTAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 8661 TGTAAAGAACTGGAGAGCCTTTAGTCGAGCGGAAACAGACCCTGTAATAATATAC 8720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 8721 AAAGAGTATCAAGAGCACTTGGTCAGAGAAATTTGGAAATTTGGAGATTAATTGAAAC 8780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 AATTAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 8781 GATTGCAAGAGATTAATCAATTGTATGAGTCAACGAGACGGAACCAAGAGAGTCTTGT 8840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 471 AGAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 8841 CGAAGAAATGTAATTTCTGTGAAAATTAACAATTCAGGGGAAACAAATCATTCATGTA 8900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 531 GAGAAAAGAAACAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 8901 AGATTAAGCCGAATTTGTGGAACCGGAAGTGAATGTGATGCTTTGATGACTTGA 8960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 591 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 8961 TGTAGATCTACAGAGAA--AAGTGTGCTATGTGTGAATTTGGTAATGTTGAT 9018
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 651 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9019 GTAATTAATGTGATTAAGAAAATGTGAAGCCAAATTCGGAAGACATGCTATTGTGAA 9078
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 711 AGAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9079 ACTACTGGAAGCAAGATTTCAAGCACTTTCAGAGCAAAAGTAATTTCTTGAAAA 9138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 771 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9139 GAGAGCGCTGACGTGAGTGAATGTGAATGTCAATTCATGATGATGATAT 9198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 831 AGAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 890
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9199 AAAAAATTTGCAAGCAAGTATGCTCTTCAATTAATAACAATAATTCAAAAATTAAT 9258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 891 AGAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 950
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9259 ATTAATTTAGCATCAAGTGAATGCAATGTGAATGTTGATGGGTGCTGAAGCAGA 9318
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 951 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 1010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9319 GGTGCAAGACCAAGGAGCACTTAATGTGAAGATTAATAATAATTAATTAATGTT 9378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1011 ACAGTAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 1070
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9379 GATTAGCAGAAAATTAATAAGAGGAGAAACA--TCAATGTATATGCGGATATGATA 9436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1071 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 1130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9437 AAAATTAATTAATTAAGTGAACAATTTCTAAGCTATTGGGATGCCAAAGTCATGCTG 9496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1131 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 1190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9497 CACTGCTGGGCACTGCCATTAATGAAAATTAAGAAATTAATTAATTAATGCGATCC 9556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 1250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9557 GAGAAATTTAAATTAATCTGGCAAGATTGGAAGGAAAGCTTAATTAATAATTCGCTGAG 9616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1251 GACAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 1310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 9617 GATCTAATCAAGTGAAGTATACGATTAATTAATTAATTAATTAATTAATTAATTAAT 9676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1311 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 1370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy 9677 AAAAAAATTCATATCAATCAAGAGAGAGAGAAAAGAAAATGTA 9726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1371 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 1420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 22
DRI25630
LOCUS
DEFINITION
DRI25630 1808 bp mRNA linear EST 15-JUN-2005
4906520 Drosophila pseudoobscura embryonic cDNA library Drosophila
pseudoobscura cDNA clone O23 5', mRNA sequence.
DRI25630
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila pseudoobscura
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1808)
REFERENCE
AUTHORS
Richards,S., Liu,Y., Bettencourt,B.R., Hradecky,P., Letovsky,S.,
Nielsen,R., Thornton,K., Hubisz,M.J., Chen,R., Meisel,R.P.,
Couronne,O., Hua,S., Smith,M.A., Zhang,P., Liu,J., Bussemaker,H.J.,
van Batenburg,M.F., Howells,S.B., Scherer,S.B., Sodergren,B.,
Mathews,B.B., Crosby,M.A., Schroeder,A.J., Ortiz-Barrientos,D.,
Rives,C.M., Metzker,M.L., Muzny,D.M., Scott,G., Steffen,D.,
Wheeler,D.A., Worley,K.C., Havlak,P., Durbin,K.J., Egan,A.,
Gill,R., Hume,J., Morgan,M.B., Miner,G., Hamilton,C., Huang,Y.,
Waldron,L., Verdusco,D., Clerc-Blankenburg,K.P., Dubchak,L.,
Noor,M.A., Anderson,W., White,K.P., Clark,A.G., Schaeffer,S.W.,
Gelbart,W., Weinstock,G.M. and Gibbs,R.A.
Comparative genome sequencing of Drosophila pseudoobscura:
chromosomal, gene, and cis-element evolution
Genome Res. 15 (1), 1-18 (2005)
15632085
JOURNAL
PUBMED
COMMENT
Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephen@bcm.tmc.edu
NCBI Trace Archive: 22697953
Insert Length: 1750 Std Error: 0.25.
Location/Qualifiers
1..1808
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/dev_stage="0-18h embryos"
/clone_lib="Drosophila pseudoobscura embryonic cDNA
library"
/notes="Vector: pOTB7; Site_1: EcoRI; site_2: XhoI; oligo
dc priming from poly A+ RNA, directionally cloned"

ORIGIN
Query Match 0.9%; Score 91.8; DB 8; Length 1808;
Best Local Similarity 42.7%; Pred. No. 1.5e-07;
Matches 557; Conservative 0; Mismatches 740; Indels 6; Gaps 2;

Qy 8421 AAGCAAGATTATGAAATATGATTAACAATAATAATTAATTCAGAGTCAATGCTCT 8480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 8481 TGCTTAATGATACAAAGATGACGATATAGATCTTTAGCGGTGCCGTGTGCA 8540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 8541 TGACAAAGAAACAACAAGCATTTACGATCAACAAGTCTTACCAACTGTAAA 8600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 AAGAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 8601 TGGAGAAACGTATCTCACTTCGTCGAAAGCTTTGGCTAAATAATTATGAAA 8660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Accession	Source	Organism	Version	Keywords	Accession	Source	Organism	Version	Keywords
9201	AAATTTGGCAGACGATGCACTCTTCATTTATATACCAAAATTCAGAAAATTAATAT	Xenopus tropicalis	1829 bp	DNA	9201	AAATTTGGCAGACGATGCACTCTTCATTTATATACCAAAATTCAGAAAATTAATAT	Xenopus tropicalis	1829 bp	DNA
610	AA	Xenopus tropicalis	1829 bp	DNA	610	AA	Xenopus tropicalis	1829 bp	DNA
9261	TACTTTAGCATCAAGTACTGATGCAATGTGATGTTGTCATGGGTGGCTGAAGCAGAGC	Xenopus tropicalis	1829 bp	DNA	9261	TACTTTAGCATCAAGTACTGATGCAATGTGATGTTGTCATGGGTGGCTGAAGCAGAGC	Xenopus tropicalis	1829 bp	DNA
580	AA	Xenopus tropicalis	1829 bp	DNA	580	AA	Xenopus tropicalis	1829 bp	DNA
9321	TGCAGAGCCCAAGGACAGTTAGTGTAAAGATCAATTAATAGACTTAATATGTTGA	Xenopus tropicalis	1829 bp	DNA	9321	TGCAGAGCCCAAGGACAGTTAGTGTAAAGATCAATTAATAGACTTAATATGTTGA	Xenopus tropicalis	1829 bp	DNA
520	AA	Xenopus tropicalis	1829 bp	DNA	520	AA	Xenopus tropicalis	1829 bp	DNA
9381	TTTAGCAGGAAATTTAAACAGAGGAAACATCAATGTATATGCGGATGTGATTAATA	Xenopus tropicalis	1829 bp	DNA	9381	TTTAGCAGGAAATTTAAACAGAGGAAACATCAATGTATATGCGGATGTGATTAATA	Xenopus tropicalis	1829 bp	DNA
460	AA	Xenopus tropicalis	1829 bp	DNA	460	AA	Xenopus tropicalis	1829 bp	DNA
9441	TTATATATTAAGTAAGACAAATTTCTAGGCTATTGGCGATGCCAAAGCTACTGTCAGC	Xenopus tropicalis	1829 bp	DNA	9441	TTATATATTAAGTAAGACAAATTTCTAGGCTATTGGCGATGCCAAAGCTACTGTCAGC	Xenopus tropicalis	1829 bp	DNA
400	AA	Xenopus tropicalis	1829 bp	DNA	400	AA	Xenopus tropicalis	1829 bp	DNA
9501	TGCTTCGGCACTGCCACTATTGAAAAAATGAACTAAATTTATATATGCGATCCGGA	Xenopus tropicalis	1829 bp	DNA	9501	TGCTTCGGCACTGCCACTATTGAAAAAATGAACTAAATTTATATATGCGATCCGGA	Xenopus tropicalis	1829 bp	DNA
340	AA	Xenopus tropicalis	1829 bp	DNA	340	AA	Xenopus tropicalis	1829 bp	DNA
9561	ATTAAAAATATCTGCGACAGTTTGAAGGAAAGCTATATAAAAAGCTGCTAGCATC	Xenopus tropicalis	1829 bp	DNA	9561	ATTAAAAATATCTGCGACAGTTTGAAGGAAAGCTATATAAAAAGCTGCTAGCATC	Xenopus tropicalis	1829 bp	DNA
280	AA	Xenopus tropicalis	1829 bp	DNA	280	AA	Xenopus tropicalis	1829 bp	DNA
9621	TATATGAGTACGCTGTATACGATTAATATATATATATATATATATATATATATAT	Xenopus tropicalis	1829 bp	DNA	9621	TATATGAGTACGCTGTATACGATTAATATATATATATATATATATATATATATAT	Xenopus tropicalis	1829 bp	DNA
220	AA	Xenopus tropicalis	1829 bp	DNA	220	AA	Xenopus tropicalis	1829 bp	DNA
9681	AAATTTGACATATCAATCAAGAGCGAGAAAGCGAAAAA	Xenopus tropicalis	1829 bp	DNA	9681	AAATTTGACATATCAATCAAGAGCGAGAAAGCGAAAAA	Xenopus tropicalis	1829 bp	DNA
160	AA	Xenopus tropicalis	1829 bp	DNA	160	AA	Xenopus tropicalis	1829 bp	DNA

ORIGIN	Library Segment 1"
Query Match	0.94; Score 91.4; DB 10; Length 1829;
Best Local Similarity	42.4%; Pred. No. 1.9e-07;
Matches 568; Conservative	0; Mismatches 763; Indels 10; Gaps 1;
QY	8383 ACACCAAAAGTATCTGAAATCCGTAGTTCGTTACGAAAGCAAGATTAGAAAATAT 8442
DB	80 AGAAGAGAAAAAACAAG 139
QY	8443 GATTACACAAAAATATATTTCTGAAAGTCATGCTCTTGCTTTAAATGATACAAAGAT 8502
DB	140 ATAAATACAAAG 199
QY	8503 GAACGGAATATAGAACTTTAGCCGTAGCCGTGTCATGCACAGAGACAAACAAGCA 8562
DB	200 AAAAAAAGATAGAACCAACAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 259
QY	8563 TTTACGAGATCAACAGTTACTTCTACACTGTAAATGAGAGAAAGTATCTCACTT 8622
DB	260 AAAAAAAG 319
QY	8623 CGTCACAAAGCTTGGCTAAAAATGAAATATATGAAATGTAAAGAACTGAGAGACC 8682
DB	320 AAAAAAG 379
QY	8683 TTACTCGAGCGGAGACAGACCGCTTGAAATTTATCAAAAGTACTTACAGAGCAATTG 8742
DB	380 AAAAAAAG 439
QY	8743 GTTCGAGAAATTTGGGAAATTTGGAGATTAATTAAGAACGATTGCAAGAGATTAATTCGATT 8802
DB	440 AAAAAAAG 499
QY	8803 GTAAGAGTCACGAGACGAAACCAAGAGAGCTTTGTCGAGAGAAATGTATTTCTGTG 8862
DB	500 GAAAG 559
QY	8863 AAAATACAAATTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8922
DB	560 AAAAAAAG 619
QY	8923 ACCGAGAGTAAATGTAGATGCTTTGATATGAATCTTGATGTAGATCTACAGAGAGAGAG 8982
DB	620 AAAAAAAG 679
QY	8983 GGTGCTATGTGGAATTTGTATTTGAAATTTGATGTAAATTAATGATTAAAGAAAT 9042
DB	680 AAG 739
QY	9043 GTAGAGCCAAATTCGAGAGACATGCTATTGTAGAACTACTGAGAGAGAGAGATATCA 9102
DB	740 GAAAG 799
QY	9103 GCATTTACAGAGCAAAATGTAATATCTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 9162
DB	800 AACAAG 849
QY	9163 ATATCGAATGTACATTTCCATGAGATGGAATTTAAATTTTGCAAGACAGATATGCA 9222
DB	850 GAAAG 909
QY	9223 TCTTCTCAATTAATACCAAAATTTCAAAATATATATTTACTTTAGCATCAAGTATGAA 9282
DB	910 AAAAAAG 969
QY	9283 TCGATGTGAATGTTCAATGGGTGAGCTGAGCAAGAGGTGACAGAGCCAAAGCAGACTT 9342
DB	970 AAAAAAAG 1029
QY	9343 AGGTAAAGATCAATTAATAGAACTTAATTTGATTTTACGAGAGAGAGAGAGAGAGAG 9402

[illegible]

D	b		1919	AAAAAAAAATTAATAATTTAAAAATTAAATTAATCAAGATAATTAAACAAAAGAAAAAATTTAA	1865
Q	y		8306	AAGAATAAGTAACTGCTAAATCTTCGTAGATGCAGAAATGAGAGAAATTGCCCGAG	8365
D	b		1859	ATAAATTAATAATAAATATATAAATCATTAATATATAAATAAATAAATTAAATTAATAA	1800
Q	y		8366	CAGAGTGGAATCTTCAACGACAAGTAATACTGAATCCGTGTTGGTTACGAAGC	8423
D	b		1799	AATATATTAATAAATAATAATAAACCAAAAAAAATTAATTAATTAATAAATAAAT	1740
Q	y		8426	AAGATTGTGAATAATATGATTACCAAAAAAATATATTTCGAAAGTCAATGCTGTGCTT	8485
D	b		1739	ATTAATAAATAAATAAATAAATTTAAATATTAACATACCTTAATAAATAAATAAATA	1680
Q	y		8486	TAAATGATACAAAGAAATGAAGCGAATATAGATCTTTAGCGGTAGCCGATGCAATCAC	8545
D	b		1679	TAAAAAATCTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1620
Q	y		8546	AAGAACAAACAAAGCAATTACGAGATCAACAAAGTTAACTTACACATGTAATGAG	8605
D	b		1619	AATTAATATCAATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1560
Q	y		8606	GAAACGTATCTCACTTCGTCAAACCTTTGCTAATAATGAATAATTATGGAATGTAA	8665
D	b		1559	AAAAATATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1500
Q	y		8666	AAGGAATGAGAGACCTTAGTCGAGCCGAAACAGCAGCCCTGAAATAATATACAAAGA	8725
D	b		1499	CAATCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1440
Q	y		8726	GTACTACAGAGCAATGTTGTCAGAAATTTGGAAATTTGAGATTAATTAGAAACGATTG	8785
D	b		1439	AAATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1380
Q	y		8786	CAAGAGTATATACGATTGTAAAGTCAACGAGACGAAACCAAGAGCTTTGTCGGA	8845
D	b		1379	AAATTAATTAATTAATTAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAA	1320
Q	y		8846	AGATGGTATTTCTGTAAAAATCAATTTCAAGGGAAACAAATTCATCCATTGAAGATA	8905
D	b		1319	AAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1260
Q	y		8906	AAGCCAGAAATGTTGGAACCGAAGTGTAATGTAGATGCTTGAATGAACCTTGATGAG	8965
D	b		1259	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1200
Q	y		8966	ATCTACAGGAAAAAGTGTGGCTATGTGTGAATTGTAATTTGCAATGTTGATGTAATA	9025
D	b		1200	-----AAAAATTAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1147
Q	y		9026	ATGATTAAGAAAAATGTGAAGCCAAATTCGAGACATGCTATGTGAACCTACTG	9085
D	b		1146	ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1087
Q	y		9086	GAATAACAAGATATCAAGCAATTTACAGAGCAAAAGTAAATATCTTGGAATAAGAGAGC	9145
D	b		1086	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1027
Q	y		9146	CTGCACTGCACTGCATATCGAATGTACATTTCCAATGAGATGATTAATAATTT	9205
D	b		1026	AATTA--TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	969
Q	y		9206	TGCGAAGCAATGCACTTCTCAATTAATTAACCAAAATTCAAAAATAATTTCTT	9265
D	b		968	TATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	909
Q	y		9266	TAGATCAAGTATGATCAATGTGAATGCTTACGCGGTGCTGAAGCAAGAGCTGAC	9325
D	b		908	AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	849
Q	y		9326	GAGCAAGGACGTTAGTGTAAAGATCAATAAATAGACTAATTAATGTTGATTGAG	9385

[illegible]

RESULT 27	LOCUS	DEFINITION
AG346189/c	1507 bp	DNA
AG346189		linear
		GSS 18-DEC-2004
		Mus musculus molossinus DNA, clone:MM901-141E08.TU, genomic survey

ACCESSION	AG346189
VERSION	AG346189.1
KEYWORDS	GSS.
SOURCE	Mus musculus molossinus
ORGANISM	Mus musculus molossinus (Japanese wild mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
Abé, K., Noguchi, H., Tagawa, K., Yuzurika, M., Toyoda, A., Kojima, T.,
Ezawa, K., Satou, N., Hattori, M., Sakaki, Y., Motiwaki, K. and
Shirotani, T.
TITLE
Contribution of Asian mouse subspecies *Mus musculus molossinus* to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
JOURNAL

PUBLISHED 155/4823
 REFERENCE 2 (bases 1 to 1507)
 AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (17-NOV-2003) Maehara Hattori, The Institute of Physical

COMMENT

library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

Sequencing : TJ
LIBRARY

```

Vector      : pBACe3.6
R.Site 1    : EcoRI
R.Site 2    : EcoRI.

```

FEATURES	Location/Qualifiers
SOURCE	1. 1507
	/organism="Mus musculus molossinus"

```

/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMG01-141E08.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMG01 Mouse Male BAC Library"

```

Query Match	0.9%;	Score 90.6;	DB 10;	Length 1507;
Best Local Similarity	38.1%;	Pred. No. 2.6e-07;		
Matches 471; Conservative	0;	Mismatches 766;	Indels 0;	Gaps 0;

[illegible]

Db 1501 AAAAAAAAAAAAAAAAAAANANANANANANANANANANANANAAAAA 1442

QY 8547 AGGACAAACAAAGCATTTACAGATCAACACAGTTAACTTCAACTGTAAATGGAGG 8606

8607 AACGATATCAACTTCGTGCAAAAGCTTTGGCTAAAAATGAAATTATGGAAATGTAAA 8666

D_b

1381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1322

[illegible]

8727 TACTACAGAGCATTGGTTCACAGAAATTGGAAATTGGAGATTAATTAGAAAAGATTGC 8786

[illegible][illegible]

8847 GAATGATTTCTGTGAAAAATACAAATTTCAGGGGAACAAATCATCCATTGAAGATA 8906

Ddb		
1141	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1082
8907	TGCGGCATTTGGTTCGAACCCGCAAGCTGCAATTGCAATGCCGTTCATAACAACCTCCTCATCCTCACACTGTAATGCTGCA	8966

Db

8967 TCTACAGAAAAAGTGTGCTACTGTCGCAATTGGTATTTGGAATGCTTAATGTAATAA 9026

Dc		9027 TGTGATTAAGAAAAATTGGTAAAGCCCAATACTCGAAGAACAACACTGCATTTGTAGAAACTGA CTGG	9086
Db	1021 NNAAAAAAAAAAAAAAAANNANNAAAAAAAAAAAAAAAAANNAANAIAAAAAAAAA	962	

Db 961 AA 902

9087 AAAACAGATATCAAGCATTTACAAGACCAAAAGTAATATTTCTTGAAGAAGACGCG 9146

9147	TGACGTGCAGCTGCATATCGAATGTACACATTCCCAATGAGATGGATAATTAAAAATTT	3206
Db	901 ANNNNAANANANANANANANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	842

Db
841 AA 782

9207 GCGAAGCAGTATGCATCTTCAATTATAACCAAAAATTCAAAAATAATATTACTTT 9266

9267	AGCATCAAGTAGTGAATCGAATGTGTCATGGGGTGGCTGAAGCAGAGCTCAGG	9326
Db	781	AA
Ox	9267	AGCATCAAGTAGTGAATCGAATGTGTCATGGGGTGGCTGAAGCAGAGCTCAGG

Db 721 AA 662

9327 AGCCAAAGCGACAGTTAGTGTAAAGAAATCAAAATAAATAGAACTAATATGTTGATTAGC 9386

D6

661 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 602

OY		9387	AGGAAAAATTTAAACAGGGGAAAATCATCAATTGTATGC GGATATGTATAAATTTTAA	9446
Dd		601	AA	542
Oy		9447	TATTAAGAACAATTTCTAAGCGCTATTGCCGAATCCAAAAGTCAGTGCTGCCTC	9506
Dd		541	AA	482
Oy		9507	GCCAATCGCCACTATTGAAAAAATGAAGTAATTTAATATCGATCCGAGAATTAA	9566
Dd		481	AA	422
Oy		9567	AATTAATCTGGCACAGATTGCAAGGAAGAACGTATATAAAAACCTCCGATGATCTATCA	9626
Dd		421	AA	362
Oy		9627	GCTAGACTGTATACGATTAATAATATACATGGCATCTTCTGAAAAAGCATCAAAAAATT	9686
Dd		361	AA	302
Oy		9687	GACATATCAATCAAAGAGAGAGAAAAAGGAAAAA	9723
Dd		301	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	265
RESULT 28 AG289820/C				
LOCUS	Mus musculus molossinus DNA,	clone:Msmg01-064824.TJ,	genomic survey sequence.	
DEFINITION				
ACCESION	AG289820	GI:47862697		
VERSION	AG289820.1			
KEYWORDS	GSS.			
SOURCE	Mus musculus molossinus (Japanese wild mouse)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus;			
REFERENCE	Abé,K., Noguchi,H., Tagawa,K., Yuzurika,M., Toyoda,A., Kojima,T., Ezawa,K., Saito,N., Hattori,M., Sakaki,Y., Moriwaiki,K. and Shiroishi,T. Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6D, as defined by BAC-end sequence-SNP analysis <i>Genome Res.</i> 14 (12), 2439-2447 (2004) 15574823			
TITLE	(bases 1 to 1401) Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki.Y. Direct Submission Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsunumi-Ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:hatterio@gsc.riken.jp, URL:http://nsp.gsc.riken.go.jp/, Tel.:81-45-503-9111, Fax:81-45-503-9170) Clones are derived from the mouse BAC library MSHG01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp			
JOURNAL	PUBMED			
REFERENCES				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
PRIMERS				
Sequencing :	TJ			
LIBRARY				
Vector :	pBac3.6			
R.Site 1 :	EcoRI			
R.Site 2 :	EcoRI			
Location/Qualifiers				
1..1401				
/organism=	"Mus musculus molossinus"			
/mol_type=	"genomic DNA"			
/sub_species=	"molossinus"			
FEATURES				
source				

Query Match	0.94; Score 90.2; DB 10; Length 1401;	
Best Local Similarity 41.8%; Pred. No. 3.1e-07;		
Matches 452; Conservative 0; Mismatches 650; Indels 0; Gaps 0;		
8641 AAAAAATGAAATTTATGCAAAATGTAAGGAACTGAGAGAGCCTTAGTCGAGCCGGAACA	8700	/db_xref="taxon:57486"
		/clone="MSMg01-064B24.TU"
1395 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGTAAATAAAAAA	1336	/sex="male"
8701 GCAGCCGCTGAAATTTATACAAAGAGTACTACAGAGACATTTGTCAGCAAAATTTGGAA	8760	/cissue type="mixture of kidney and spleen"
		/clone_lib="MSMg01 Mouse Male BAC Library"
1335 AAGAAAAAAB	1276	
8761 ATTGAGATTAATTAGAAACGATTGCAGAGATATATCGATTGTAAAGTCAACGAGAC	8820	
1275 TMAAAAAAAAAAAAAAGAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAAT	1216	
8821 GGAACCAAGAGAGCTTGTCCGAAAGAAATGCTATTTCTGGAATAATCAATTTCACGG	8880	
1215 AAAAAAAAAAAAAAAAAAAAAA CAAAAACAAAAAGAAAAAGAAAAAGAAAAAGAGAA	1156	
8881 GAAACAAAAATCATCTCATTTGAAATPAAAGCCGAAATTTGTCGACCGAAGTGAATGTA	8940	
1155 AAAAAATTAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	1096	
8941 GATCCTTGAATGAACCTTGATGATCTACAGAGAAAAAGTGTGCTATGTGGAATT	9000	
1095 CAAAAAAAAAAAAAAAAAAAAA GAAAAAAAAAAAAAAAAAAAAAGAGAAAAAGAAAAA	1036	
9001 GGTATTGGAATGTTGATGTAAATATGTGATTAAAGAAAAATGTAGAACCAAAATCGGA	9060	
1035 AGGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	976	
9061 AGACATGCTATTGTAGAAACTACTGAAAAACAAGATATCAAGCATTTTCAAGAGCAAA	9120	
975 AAAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	916	
9121 GTAAATATTCTTGGAAAAAGAGACGCTGACAGCTGACATGCAATGTAACATT	9180	
915 GAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	856	
9181 TCCATGATGATATTTAAAAATTTGGCAAGCAGTATGCACTTCTCAATTATTAAC	9240	
855 AAAAAAAAAAAAAAAAAAAAAA GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA	796	
9241 AAAAAATCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	9300	
795 AAAAAAGAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA	736	
9301 GGGGTGCTGAAAGCAGAGGTGTCAGAGCCAAAGCAGTATAGTGAAGATCAATA	9360	
735 AAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA	676	
9361 AATGAACATAATATGTTGATTAGCAGAAAAATTTAAACAGAGGAAACATCAATGTA	9420	
675 AAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA	616	
9421 TATCCGCGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	9480	
615 AAAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	556	
9481 GCCAAAAATCATGCTGAGCTCTTCGCAACTGCACTATTGAAAAAATGAAGTAAA	9540	
555 GAAAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	496	
9541 TTTAATAATGCGATCCGAAATTTAAAAATATATGTCGCAAGATTGAAAGGAAAGCTAAT	9600	

Db	495	NNNNNNNTTGGCGGCGGGAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAA	438		
OY	9601	AAAAAACAGTCGTGGATCTTAATCAGTAGACTGCCTATTACGATAAATATCATGCGAT		9660		
Db	435	AAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAA	376		
OY	9661	TCTTCTGAAAAAGCATTACAAAAATTTGATATCAATCAAAGAGCGAAAAAGGAAA		9720		
Db	375	AAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAA	316		
OY	9721	AA	9722			
Db	315	AA	314			
RESULT 29		AJ592026	1407 bp	DNA	linear	GSS 15-JAN-2004
LOCUS		AJ592026/c				
DEFINITION		Arabidopsis thaliana T-DNA flanking sequence, right border, clone 600D05, genomic survey sequence.				
ACCESSION		AJ592026				
VERSION		AJ592026.1	G1:3794350			
KEYWORDS		GSS; right border; T-DNA flanking sequence.				
SOURCE		Arabidopsis thaliana (chale cress)				
ORGANISM		Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE		1	Brunaud V., Balzerque S., Dubreucq B., Aubourg S., Samson F., Chauvin S., Bechtold N., Cnuad C., DeRose R., Pelletier G., Lepiniec L., Caboche M. and Leclercy A. T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites EMBO Rep. 3 (12), 1152-1157 (2002) 1244555 2 (bases 1 to 1407) Balzerque S. Direct Submision Submitted (23-OCT-2003) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE PCR was performed on DNA from transformants of Arabidopsis thalianana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at http://dbgap.versailles.inra.fr/publiclines/. This sequence has been generated in the framework of the French plant genomics program "Genoplante" (http://www.genoplante.com) and http://genoplante-info.infobiogen.fr/ .			
COMMENT		Location/Qualifiers				
FEATURES		source				
		1..1407	/organism="Arabidopsis thaliana"			
			/mol_type="Genomic DNA"			
			/db_xref="taxon:3702"			
			/clone="600D05"			
			/clone_1lb="Arabidopsis thaliana T-DNA insertion lines"			
			/ecotype="Masseleskija"			
			1..1407			
			/note="T-DNA flanking sequence			
			right border"			
ORIGIN						
OY	8421	AAAGCAAGTTATGAAATTAAGATTCACAAAAAATATATTTCAGAGCATGCTCT				8480
Db	1330	AAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAA			1271
		Query Match	0.9%;	Score 90.2;	DB 10;	Length 1407;
		Best Local Similarity	37.5%;	Pred. No. 3.1e-07;		
		Matches 488;	Conservative	0;	Mismatches 815;	Indels 0;
				Gaps	0;	

QY	8481	IGCTTTAAATGATCAAAAGATGAGCGAATATGAAATCTTTACCGGTGCCGCGTGTGCA	8540
DB	1270	AA	1211
QY	8541	TGCACGAAGACAACAAGACTTATCGAGATCAACAAGTTAACTTCAACACTGTAAA	8600
DB	1210	AA	1151
QY	8601	TGGAGAAACGTATCTCAACTTCGTGCAAAAGCTTTGGCTTAATAATGAAATTAATGAAA	8660
DB	1150	AA	1091
QY	8661	TGTAAAGAACTGAGAGACCTTAGTCGAGCGGAAACAGCACCGCTGTAAATTTATAC	8720
DB	1090	AA	1031
QY	8721	AAAGAGTACTACAGAGACATGTTGTTGACGAAATTTGGAAATTTGAGATAAATTAGAAAC	8780
DB	1030	AA	971
QY	8781	GATTGACAGAGATATATCGATTGTAAAGTCAACGAGACGGAACCAAGAGAGTCTTGT	8840
DB	970	AA	911
QY	8841	CGGAAGAAATGTATTTCTGTGAAATATCAATTTACGGGAAACAATATCATCATTTGA	8900
DB	910	AA	851
QY	8901	AGATTAAGCCGAATTTGTGGAACCGGAAGTGTAAATGTAGATCTTTGAATGAACCTTGA	8960
DB	850	AA	791
QY	8961	TGTAGATCTACAGAAAAAGTGTGCTATGTGGAATTTGTAATGTTGATGT	9020
DB	790	AA	731
QY	9021	AAATATATGTGTTTAAGAAAAATGTAAAGCCAAATTCGAAAGATGCTATTTTGAAC	9080
DB	730	AA	671
QY	9081	TACTGAAAAACAATATATCAAGACTTTACAGAGCAAAAGTAAATTTCTTGGAAAAAG	9140
DB	670	AA	611
QY	9141	AGACGCTGCAGCTGCAGCTGCATATCGAATGTACACATTTCCAATGAGATGATATTA	9200
DB	610	AA	551
QY	9201	AAATTTGGCAAGCAGATATGATCTTCTCAATTAATCACAAATTCCAAAAAATTAATAT	9260
DB	550	AA	491
QY	9261	TACTTTAGCATCAATGATGATCGAATGTGAATGTTCAATGGGTGGCTGAAGCAAGAG	9320
DB	490	AA	431
QY	9321	TGCAGAGCCAAAGCGACAGTTAGTGAAGATCAATAATATGAACCTTAATTAATGTTGA	9380
DB	430	AA	371
QY	9381	TTTTCGACGAAAAATTAAACAGAGGAAACATCAATGTATATCCGGAATTTGATTAATA	9440
DB	370	AA	311
QY	9441	TTATATATATAGTAAGCAAAATCTTAAGGCTATTGCGAGTCGCAAAAGTCAATGCTGCAC	9500
DB	310	AA	251
QY	9501	TGCTTCGCAACTGCCATATTTGAATAAAATGAAGTAAATTTATATATGCAATCGAGA	9560
DB	250	AA	191
QY	9561	ATTTAAAAATATCTGCAGAGATTGGAAGGAAAGCTAATAAAAAAACGTCGATGAGATC	9620

[illegible]

LOCUS	DEFINITION	ACCSSION	VERSION	KEYWORDS	SOURCE	ORGANISM
CL027236	1618 bp DNA linear					
CL027236	CH216-25H10_Rm1.1 CH216 <i>Xenopus tropicalis</i> genomic clone					
CL027236	CH216-6SH10, genomic survey sequence.					
CL027236	CL027236					
CL027236.1	GI:40471752					
GSS.						
<i>Xenopus tropicalis</i> (western clawed frog)						
<i>Xenopus tropicalis</i>						

REFERENCE	1 (bases 1 to 1618)
AUTHORS	Kremiczki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE	A physical map of the xenopus tropicalis genome
JOURNAL	Unpublished (2003)
COMMENT	Contact: Richard K Wilson

Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert length: 175000 Std Error: 0.00
Seq primer: RML TACGACTCTACTATGGGAGA
Class: BAC ends
High quality sequence start: 1281
High quality sequence stop: 1335.

```

FEATURES
source
location/Qualifiers
1..1618
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-25H10"
/sex="male"
/cell_line="Strock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: pTARBAC2.1; CHO1-216 Xenopus tropicalis
BAC library"

```

[illegible]

Qy		8881	GAAACAAAATCATCTGTAAGATTAAAGCCGAATTTGTGGAACCGGAAGTGTAAATGTA	8940
Db		590	GAGAGAAAAAGAGCAGAAGAGAAAGAAAGAGANANANANAGAGAAAAAAGAGCAATA	649
Qy		8941	GATGCTTTGATGAACTTGATGTAGATCTCAAGAGAAAAAGTGTGCCTATGTGTGAATT	9000
Db		650	ATTATTTATTAATGTAAAGAAAAAABAGA - AAGAGAAAAAAGAGAGCGGGAGGA	708
Qy		9001	GGTATTGGAATGTGTA TGTAAATATGTGATTAGAAAAATGTAGAAACCBAATCGGA	9060
Db		709	AGAAAAAAGAAAAAGAAAAAAGTGGAGAAAAAAGAAAAAGAAATTAAGA	768
Qy		9061	AGACATCTATTGTAGAAACTACAGAAACAAGATTCAGAGCTTACAGAGCAAA	9120
Db		769	AAAAAAAAAAAAAAAAATAGAGAAATTAATAAAGAAAAAGAAAGAAAAAAGAAAAAGAA	828
Qy		9121	GTAATATTTCTGGAAAAAGAGACGCTGCAGCTGCAGCTGCAGATATCGAAATGTACACTT	9180
Db		829	AAAAAAAAAAAAAAAAATTAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	888
Qy		9181	TCCAATGAGATGATATTAATAATTTGGCAAAGCAGTAGTCATCTCTCAATTATTAACC	9240
Db		889	AAAAAGAGAAAAAAGAAAAAAGAAAAAATGAATAAAGAAAAAAGAAAAAAGAAAAA	948
Qy		9241	AAAAATTCAAAAAATAATATTTACTTTGCAATCAAGTAGTGATGTGAATGTGCAT	9300
Db		949	AAA	1008
Qy		9301	GGGGTGGCTGAGCAAGGTCGACGAGCCAAAGCGCACGTTAGTGTAAAGATCAATA	9360
Db		1009	AAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG	1068
Qy		9361	AATGAACTAATATATGTTGATTTAGCAGAAAAATTAAAA CAGAGGAAACATCAATGTA	9420
Db		1069	AAA	1128
Qy		9421	TATCCGGATATGATAAAAATTATATATTAATAGTAACTTAAGCTATTGCCGAT	9480
Db		1129	AAA	1188
Qy		9481	GCCAAAGTATGCTGAGCTGCTTCGGCAACTGCCCTATTGAAAAAATGAAGTAAA	9540
Db		1189	AAAAAAAAAAAAAAAAATTAATAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1248
Qy		9541	TTTTATATATCGAATCCGAGATTTTAATAATATCTGCAAGATTTGAAAGGAAGCTAAT	9600
Db		1249	AAA	1308
Qy		9601	AAAAAAACGTCGTAGGATCTAATCAGGTAGACTGTATACGATTAATATCATGTCAT	9660
Db		1309	AAA	1368
Qy		9661	TCTTCTGAAAAAGCATACAAAAATTGACATATCAAAAGAGAGGAGAAAAAGGAAA	9720
Db		1369	AAAAAAAAAAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG	1428
Qy		9721	AAA 9723	
Db		1429	AAA 1431	

LOCUS	CL081234	1824 bp	DNA	linear	GSS 31-DEC-2003
DEFINITION	CH216-161B9, RM4.1 CH216 <i>Xenopus tropicalis</i> genomic clone				
ACCESSION	CH216-161B9, genomic survey sequence.				
VERSION	CL081234				
KEYWORDS	CL081234.1	GI:40537147			
SOURCE	GSS.				
ORGANISM	<i>Xenopus tropicalis</i> (western clawed frog)				
	<i>Xenopus tropicalis</i>				
	Amphibia; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;				

TITLE Gelbart, W., Weinstock, G.M. and Gibbs, R.A.
Comparative genome sequencing of *Drosophila pseudoobscura*:
Chromosomal, gene, and cis-element evolution
JOURNAL Genome Res. 15 (1), 1-18 (2005)
PUBMED 15632085
COMMENT Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephen@bcm.tmc.edu
NCBI Trace Archive: 226717045
Insert Length: 1750 Std Error: 0.25.

FEATURES
source location/Qualifiers
1..1932
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="113"
/dev_stage="0-18h embryos"
/clone_lib="Drosophila pseudoobscura embryonic cDNA
library"
/note="Vector: pOT87; Site 1: EcoRI; site 2: XhoI; oligo
dc priming from poly A+ RNA, directionally cloned"

ORIGIN

Query Match 0.9%; Score 89.2; DB 8; Length 1932;
Best Local Similarity 41.4%; Pred. No. 5.3e-07;
Matches 538; Conservative 0; Mismatches 763; Indels 0; Gaps 0;

8422 AAGCAGATTATGAAATATGATTCACAAAAAATATTTTTCAGAGTCATGCTT 8481
264 AAAN 323
8482 GCTTAAATGATCAAGATGAGCGAATATGAAATCTTTAGCGTAGCCGTGCAAT 8541
324 AAAAAACAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 383
8542 GCACAAGAACAAACAGATTTCAGATCAACAGTTAATCTTCACTGTAAAT 8601
384 AAAAAAAGAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 443
8602 GGAAGAAAGTATCTCACTGCGCAAAAGCTTTGGCTAAATGAAATTTGAAAT 8661
444 AAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 503
8662 GTTAAAGAACTGAGAGCCTTAGTCGAGCGGAAACAGACCCGTTGAAATTTATCA 8721
504 AA 563
8722 AAGAGTACTACAGACATTTGTTGCAAGAAATTTGGAAATTTGAGATTAATTGAAACG 8781
564 GAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 623
8782 ATTGCAAGATATATGATTTGATGAGTCAACGAGACGAGCAACAAAGAGCTTTGTC 8841
624 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 683
8842 GGAAGAAAGTATTTCTGTGAAAAATACATTTTCAGGGGAAACAAATCATCTATGAA 8901
684 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 743
8902 GATTAAGCCAGAAATTTGAGACCGAGAGTAAATGTAGATCTTTGAAATGAATTTGAT 8961
744 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 803
8962 GTAGATCTACAGAAAAAGTGTGCTATGTGGAATTTGTAATTTGTAATTTGTAAT 9021
804 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 863
9022 AATTAATGTGATTAAGAAAAATGTAGAACCAAAATCGAAGACATGCTATTTGTGAAC 9081
864 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 923

9082 ACTGGAAGAACAGATATACAGATTTCACAGAGCAAAAGTAAATTTCTTGGAAAAAGCA 9141
924 AAAGAACG 983
9142 GACGCTGACAGCTGACAGCTCAATATCGAATGTACATTTCCAAATGAGATATATAA 9201
984 AAGAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1043
9202 AATTGGCAAGACAGTATCATCTTCTCAATTAATACCAAAATTCAAAAATTAATAT 9261
1044 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1103
9262 ACTTTAGCATCAAGTATGTAATCGAATGCAATGTTCAAGGGGTGCTAAGCAAGAGT 9321
1104 AAAGAAAA 1163
9322 GCGAGAGCCAAAGCGACAGTATGTAAGAAATCAATTAATTAATTAATGTTGAT 9381
1164 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1223
9382 TTACAGAGAAAAATTTAAACAGAGGAAACATCAATGTATATCCGATATGATTAAT 9441
1224 AAAGAGAAAAAAGAAAA 1283
9442 TATATATATATGTAAGCAAAATTTAGGCTATTTCCGAGCCAAAGTCATGCTGACGT 9501
1284 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1343
9502 GCTTGGCAACTGCCATTTGAAAAAATGAAATTTTAATTAATGATCCAGAA 9561
1344 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1403
9562 TTTAAATATATCTGGCAAGATTGAGGAAAGCTAATTAATAAGCTCGTAGATCT 9621
1404 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1463
9622 AATCAGTAGACTGTATGCGATTAATTAATCAATGCTTTCTGAAAAAGCATCAAA 9681
1464 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1523
9682 AAATGACATATCAATCAAGAGAGAGAAAAAGGAAAAA 9722
1524 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1564

RESULT 33
CC187638
LOCUS 1434 bp DNA linear GSS 08-MAY-2003
DEFINITION CH261-98P15_Sp6.1 CH261 Gallus gallus genomic clone CH261-98P15,
genomic survey sequence.
ACCESSION CC187638
VERSION CC187638.1 GI:30431538
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1434)
Krentzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardle, E. and Wilson, R.
Gallus gallus BAC End Reads
Unpublished (2003)
CONTACT: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@atson.wustl.edu
Insert Length: 18200 Std Error: 0.00
Seq primer: Sp6 ATTATGTGACACTATAG
Class: BAC ends
High quality sequence start: 422
High quality sequence stop: 487.

FEATURES Source

Location/Qualifiers
1. 1434
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-98P15"
/sex="female"
/cell_line="UCD001, Indred 256"
/clone_1ib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken Library - For library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN

```

Query Match      0 %; Score 88.6; DB 9; Length 1434;
Best Local Similarity 40.8%; Pred. No. 6.6e-07;
Matches 447; Conservative 0; Mismatches 648; Indels 1; Gaps 1;

Qy 8628 AAAAGCTTTGGCTAAATGAAATATGAAATGTAAGAACTGAGAGCTTACT 8687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 72

Qy 8688 CCGAGCGAAGACGACGCCGTTGAAATATACAAAGTACTACAGAGCATGGTTC 8747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 132

Qy 8748 AGGAATTTGGGAAATTTGAGATTAATTGAAACGATTGCAAGAGATTAATGCTAG 8807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 192

Qy 8808 AGTCAACGAGACGGAACCAAGAGAGTCTTGTCCGAAAGATGTAATTTCTGTGAAA 8867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 252

Qy 8868 TACAAATTTGAGGGAAACAAATCATCATTTGAAGTAAAGCCAGATTGTTGAAACCG 8927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 312

Qy 8928 AAGCTAAATGTGATGCTTTGAACTGATGATGATGATGATGATGATGATGATGATG 8987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 372

Qy 8988 CTATGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 432

Qy 9048 AGCCAAATTCGAGAGCATGTAATGTAGAACTAGTGAACCAAGATATCAAGCATT 9107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 492

Qy 9108 TACAAAGAGCAAAAGTAATTTCTTGGAAAAGAGCGCTGCAAGCTGCAATATTC 9167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 552

Qy 9168 GAATGACATTTCCATGAGATGATATTAATAATTTGCAAGACAGATGATGATTC 9227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 CCCCCCCNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 612

Qy 9228 TCAATTAATACCAAAATTTCAAAATATATATTACTTACATCAAGTAGTGAATGAA 9287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCCCCCAAAAAAAAAA 672

Qy 9288 TGTGATGTTTCATGGGGTGGCTGAAGCAGAGGTGCAAGAGCCAAAGCAGTATGTGT 9347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 673 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 732

Qy 9348 AAAAGATCAATTAATAGAACTAATATGTTGATTTAGCGAGAAATTAATAACAGAGG 9407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 733 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 792

Qy 9408 AAACATCAATGATATGCGGATATGATTAATAATTAATATAGTACAAATTTCTAA 9467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 793 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 852

Qy 9468 GCGCTATTGCGGATGCCAAAAGTCATGCTGAGCTGCTTCCGGAACCTGACATTTGAAA 9527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 853 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 912

Qy 9528 AAATGAGTAAATTAATTAATGATCGATCGAGAAATTTAAATAATTCGCAAGATTGGA 9587
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 913 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 972

Qy 9588 AGGAAAGCTAATAAAACCTCGTGAATCTAATCAGTAGCTGTGATACGATTA 9647
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 973 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1031

Qy 9648 ATATACATGCGCATCTTCTGAAAAAGCATACAAAAATTTGACATATCAATCAAGAGG 9707
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1032 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1091

Qy 9708 AGAAAAAGGAAAAAA 9723

Db 1092 AAAAAAAAAAAAAAAAAA 1107

```

```

RESULT 34
DR147304
LOCUS
DEFINITION
Drosophila pseudoobscura embryonic cDNA library Drosophila
pseudobscura cDNA clone O11 3', mRNA sequence.
ACCESSION
DR147304
VERSION
DR147304.1 GI:67892394
KEYWORDS
EST.
SOURCE
Drosophila pseudoobscura
Drosophila pseudoobscura
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 2071)
Richards, S., Liu, Y., Bettencourt, B.R., Hradecky, P., Letovsky, S.,
Nielsen, R., Thornton, K., Hubisz, M.J., Chen, R., Weisell, R.P.,
Coutome, O., Hua, S., Smith, M.A., Zhang, P., Liu, J., Buesemeyer, H.J.,
van Batenburg, M.F., Howells, S.L., Scherer, S.B., Soderstrom, B.,
Matthews, B.B., Crosby, M.A., Schroeder, A.J., Ortiz-Barrientos, D.,
Rives, C.M., Metzker, M.L., Muzny, D.M., Scott, G., Steffen, D.,
Wheeler, D.A., Worley, K.C., Havlak, P., Durbin, K.J., Egan, A.,
Gill, R., Hume, J., Morgan, M.B., Miller, G., Hamilton, C., Huang, Y.,
Malden, L., Verdusco, D., Clerc-Blinkenburg, K.P., Dubchak, I.,
Noor, M.A., Anderson, M., White, K.P., Clark, A.G., Schaeffer, S.W.,
Gelbart, W., Weinstein, G.M., and Gibbs, R.A.
Comparative genome sequencing of Drosophila pseudoobscura:
chromosomal, gene, and cis-element evolution
Genome Res. 15 (1), 1-18 (2005)
15632085
JOURNAL
PUBMED
COMMENT
Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephenr@bcm.tmc.edu
NCBI Trace Archive: 226691688
Insert Length: 1750 Std Error: 0.25.
Location/Qualifiers
1. 2071
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="O11"
/dev_stage="0-18h embryo"
/clone_1ib="Drosophila pseudoobscura embryonic cDNA
library"
/note="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo
dt priming from poly A+ RNA, directionally cloned"

```

FEATURES source

1. 2071
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="O11"
/dev_stage="0-18h embryo"
/clone_1ib="Drosophila pseudoobscura embryonic cDNA
library"
/note="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo
dt priming from poly A+ RNA, directionally cloned"

ORIGIN

Query Match	Similarity	0.9%;	Score 87.8;	DB 8;	Length 2071;
Best Local Match	Similarity	41.0%;	Pred. No. 1e-06;		
Matches	Conservative	0;	Mismatches 1016;	Indels	Gaps
Oy	7110	AGAAAAAGTAAATGTATGACGCAAAAGGAGAAAAATCTTAAAGCAACAGACGAATGC	7169		
Db	252	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	311		
Oy	7170	AGAAATCGAGAGACGACGAATCGAGCCAAATGTCTTGTAATAATTTGGACAGCTGT	7229		
Db	312	AAAAAGAAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	371		
Oy	7230	AGAAATGTGAAAAAATTTCTGAAGAAAAAGAACGAGATTAAAACTTTAGACGAAGT	7289		
Db	372	GAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	431		
Oy	7290	TAAACAAAGACAAAGTAAAAAAAGTAAATGATGCTACGAAAAAAATCTTACAATCAGCAG	7349		
Db	432	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	491		
Oy	7350	TATTTCTACAGAGATCTTCTGTAAAAACGGATAGAGAGATCTCAGGAGAGAGAT	7409		
Db	492	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	551		
Oy	7410	TAAAGCCTTGTGAAGACTTCTGATATTATTGAAAAAATAGATTATTCACACAGCA	7469		
Db	552	GAAGAAAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	611		
Oy	7470	CAGAATAATATCACTTCTACGTGCTTGGGAATCGACAGTCTGCTCCGATCAG	7529		
Db	612	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	671		
Oy	7530	AACAGTCGAGTTTCAAAATATTTAAAGAAATTCGAGTTTCTGTGAAAAATTCCTTTGT	7589		
Db	672	AAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	731		
Oy	7590	GAAGCAGCTGAAAAAGTAAATGTTAGTCCGATTTACAGAAATGTTGCTTTAACAC	7649		
Db	732	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	791		
Oy	7650	ATATCAAGTCTGTAGAGCATTTGGCAATAGAGCTGCCATTCAGATTTAAATTTCTAA	7709		
Db	792	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	851		
Oy	7710	TGCAAGATCAATATATCAGTATTTAAAAATTTCTAAGCTATTAGAAAAAATATTGATGTAT	7769		
Db	852	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	911		
Oy	7770	TGTAAAAAGATTAATCGAAATTGAGACCGAAACAAAGATTAAACCTAGAGCGGTAGC	7829		
Db	912	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	971		
Oy	7830	TGCCGAGCCATTTCTCAAAAGCAAAGATGAAATGAATTCAGAGGTTGAAATTGAGAA	7889		
Db	972	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	1031		
Oy	7890	GAGTATTTTCATGAGAAAAATPAGAGTAATACGCCCTCTTAAGGAATTTGAGAGAAAT	7949		
Db	1032	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	1091		
Oy	7950	CAATGTCAAAGTGAAGAAAAAGAAACAGAGTACGTCTGAATCTCAGAGACTTCTGTAGG	8009		
Db	1092	AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	1151		
Oy	8010	AGCAGTACAGGGGCGAGATTTATTTCCGAAGCAAAAGATGCGCGAGACTTTATTGGAA	8069		
Db	1152	AAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	1211		
Oy	8070	AGTTAGTCAAAATATCCGAGAGATTTTTCATGCAATATGTAATGGAAGCAAC	8129		
Db	1212	GAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	1271		
Oy	8130	ACATTAATGAAGTAAAGCATTTCTTAAGCAGTAAACAGTCTCTGATTGGAGAGT	8189		

D	b		1272	AAAGGAAAAAAGNAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAA	1331
O	y		8190	TGGAGTCAACAGGCAGAGCTACTGTCCAGGTAAAATCATGTAGATTGAAGAAG	8249
D	b		1332	AAGAAAAAAAAAAGAAAAAAAAANAAAAAAAAAAAAAAAAAGAAAAAAAAA	1391
O	y		8250	AAATTGTTCAGAACAAATCGATTGAATTCATAATTCTAAAGTGACGTTGATGACA	8309
D	b		1392	GAAAAAAAAAAAAAAAAAGAGAGAAAAAAAAAGAAAAAAAAAGAAAAAAAAAG	1451
O	y		8310	TAAAGTAAGTCTAAATCTTCTGTATGATCAGAAATGAGAGAAATTCGCCGAGCAG	8369
D	b		1452	AAGAAAAAGAAAAAAAAAGAAAAAAAAAGAAAAAAAAAGAAAAAAAAAG	1511
O	y		8370	AGTAAATACCTTCTCACACACAAAGTAATACTGAAATCCGATGCTTTACGAAACAGA	8429
D	b		1512	AAAGAAAAAAAAAGAAAAAAAAAGAAAAAAAAAGAAAAAAAAAGAGCAA	1571
O	y		8430	TTATGAAATATATGATTATACACAAAAATATATTTCAAGATCAATGCTTCTTCTTAA	8489
D	b		1572	AAAAAGAAAAAAAAAAAAACAGAAAAAGAAAGAGAAAAA	1631
O	y		8490	TGATACAAAGATGAACCGAATATGAAATCTTAAAGCGTACCCTGTGCATGCACAG	8549
D	b		1632	AGAAAAAAAAAAAAAAAAAGAAAAAAAAAGAAAGAAAAAAAAAGACAGAAAG	1691
O	y		8550	AACAACAAAGCATTTACGAGATCAACCAAGTTACTTACAACTTAATGAGAGAA	8609
D	b		1692	AAAAAAAAAGAAAGAAACAGAAAAAGAA--AAAAAAAAAGAAAAAAAAA	1749
O	y		8610	CGATCTCAACTTCGTCCAAAGCTTTGCTAAATAATGAAAAATTAGAAATGTAAAG	8669
D	b		1750	AAAAAAAAAAAAAAAAAGAAAGAAAGAAAGAAAAAAAAAGAAAGAAAAAAAAA	1809
O	y		8670	AACGTGAGAGCCTTAGTCGAGCGGAAACAGCAGCCGTTGAATAATTATCAAGAGTAC	8729
D	b		1810	AAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAA	1869
O	y		8730	TACAGGAGCAATGTGTCCAGAAATTGCGAATTTGAGATTAATTAGAAAGATTGCAAG	8789
D	b		1870	AAAAAAAAAAAAAAAAAGAAAGAAAGAAAAAAAAAGAAAGAAAGAAAGAA	1929
O	y		8790	AGATATATCGATTGTGTAAGTCAACGAGACGAAACCAAGAGAG	8834
D	b		1930	AGAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAGAAAGAAAG	1974
RESULT 35					
CL082658 1457 bp DNA linear GSS 31-DEC-2003					
LOCUS CH216-165B13.RM4.1 CH216 Xenopus tropicalis genomic clone					
DEFINITION CH216-165B13, genomic survey sequence.					
ACCESSION CL082658					
VERSION CL082658.1 GI:40538571					
KEYWORDS GSS.					
SOURCE Xenopus tropicalis (western clawed frog)					
ORGANISM Xenopus tropicalis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipodidae; Pipidae;					
REFERENCE 1 (bases 1 to 1457) Kremetzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.,					
AUTHORS A physical map of the xenopus tropicalis genome unpublished (2003)					
TITLE Genome Sequencing Center JOURNAL Washington University School of Medicine COMMENT Email: submissions@watson.wustl.edu Insert Length: 175000 Std Error: 0.00 Seq primer: RM4 ctcaaggcatcgctcagac Class:BAC ends					

ORIGIN

/note="Vector: PTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

Query Match 0.94; Score 87.4; DB 10; Length 1428;
Best Local Similarity 37.5%; Pred. No. 1.2e-06;
Matches 496; Conservative 0; Mismatches 826; Indels 0; Gaps 0;

QY 8402 AATCCGAGTTCCTTACGAAAGCAAGATTATGAAATTAATGATTACACAAAAATATA 8461
DB 96 AA 155
QY 8462 TTTCAGAGTCAATGCTCTTGCTTTAATGATACAAAGATGAAGCATATATGATCTT 8521
DB 156 NNN 215
QY 8522 TAGCGTAGCCGCTGTCATGCAACAGAAACAAACAAAGATTACAGATCAACAAAGT 8581
DB 216 NNNNNNNNNCCNNNAA 275
QY 8582 TAACTTCTACACTGTAAATGAGAGAAACGTATCTCACTTCGTGCAAAAGCTTTGGCTA 8641
DB 276 AA 335
QY 8642 AAATGAAATTTAGAAATGTAAAGAACTGAGAGCGCTTAGTGGAGCGGAAACAG 8701
DB 336 AA 395
QY 8702 CAGCCGTTGAAATTTATACAAAGATCTACAGAGCATTTGTTGACAGAAATTTGGGAA 8761
DB 396 AA 455
QY 8762 TTGAGATTAATTAAGAACGATTCAGAGATTAATGATTTAGAGTCAACGACGACG 8821
DB 456 AA 515
QY 8822 GAACCAAGAGGCTTGTCGAAAGAAATGTATTTCTGAAATAATACATTTACAGGG 8881
DB 516 AA 575
QY 8882 AAACAAATCATTCATTGAAGATTAACCAAGATTGTTGACCGAACTGAATGTAG 8941
DB 576 AA 635
QY 8942 ATGCTTTGAATGAATCTTGATGTGATTCACAGAAAGAGTGTGCTATGTGGAATTG 9001
DB 636 AA 695
QY 9002 GTATTGGAATGTGATGTAATTAATGATTAAGAAATGTAAGACCAAAATCGGA 9061
DB 696 AA 755
QY 9062 GACATGCTATTGTGAATACTGAAACAAAGAAATTCAGCACTTACAGAGCAAA 9121
DB 756 AA 815
QY 9122 TAAATATTCTTGAGAAAGAGCGCTGACCTGCAATTCGATTCGATCACTT 9181
DB 816 AA 875
QY 9182 CCAATGATGATTAATTAATTTGGCAAGCAAGCTGATCTTCTCAATTAATCA 9241
DB 876 AA 935
QY 9242 AAATTCAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9301
DB 936 AA 995
QY 9302 GGTGCTGGAAGCAAGGCTGACAGCCAAAGCAAGCAATTAATTAATTAATTAATTA 9361
DB 996 AA 1055
QY 9362 ATGAACTAATTAATGTTGATTTAGCAGAAATTAATAAGAGGAAATCAATGAT 9421

DB 1056 AA 1115
QY 9422 ATCCCGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9481
DB 1116 AA 1175
QY 9482 CCAAAAGTCAATGCTGAGCTGCTCGCAACTGCCATTTGAAAAAATGAAGTAAAT 9541
DB 1176 AA 1235
QY 9542 TTAATAATGATCCGAGATTTAAATAATGATTCGCAAGATTGAAAGGAAGCTAATA 9601
DB 1236 AA 1295
QY 9602 AAAAAAGCTGAGATCTAATCAGTAGACTGTATACGATTAATATATACATGCGATT 9661
DB 1296 AA 1355
QY 9662 CTCTGAAAAAGCATCAAAAAATTTGACATATCATCAAGAGAGAGAAAAAGGAAAA 9721
DB 1356 AA 1415
QY 9722 AA 9723
DB 1416 AA 1417

RESULT 37
AJ927291/c 1139 bp mRNA linear EST 10-JUN-2005
LOCUS AJ927291 Theileria annulata merozoite Theileria annulata cDNA clone
DEFINITION cam028d02.gbk, mRNA sequence.
ACCESSION AJ927291
VERSION AJ927291.1 GI:67497675
KEYWORDS EST.
SOURCE Theileria annulata
ORGANISM Theileria annulata
Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileridae;
Theileria.

REFERENCE 1 (bases 1 to 1139)
AUTHORS Pain,A., Renaud,H., Berriman,M., Murphy,L., Yeats,C.A., Weir,W.,
Kerhornou,A., Aslett,M., Bishop,R., Bouchier,C., Cochet,M.,
Coulson,R.M.R., Cronin,A., de Villiers,E., Frazer,A., Foster,N.,
Gardner,M., Goble,A., Maser,P., McKellar,S., Harris,D.E., Katzer,F.,
Laird,N., Lord,A., Maser,P., McKellar,S., Mooney,P., Morton,F.,
Nene,V., O'Neill,S., Price,C., Quail,M.A., Rabinowitsch,E.,
Rawlings,N.D., Rutter,S., Saunders,D., Seeger,K., Shah,T.,
Squares,R., Squares,S., Tivey,A., Walker,A.R., Woodward,J.,
Dobbelare,D.A.E., Langsley,G., Rajandream,M.-A., McKeever,D.,
Shiels,B., Tait,A., Bartell,B. and Hall,N.
The genome of the host-cell transforming parasite Theileria
annulata and a comparison with T. parva
Unpublished (2005)

TITLE JOURNAL
COMMENT The Pathogen Sequencing Unit
The Wellcome Trust Sanger Institute
Genome Campus, CB10 1SA, UNITED KINGDOM
Merozoite cDNA library: Frank Katzer and Brian Shiels, Division of
Veterinary Infection and Immunity, ICM, University of Glasgow, UK.
Location/Qualifiers
1.1139
FEATURES
source

/organism="Theileria annulata"
/mol_type="mRNA"
/isolate="Ankara (clone D7)"
/db_xref="taxon:5874"
/clone="cam028d02.gbk"
/dev_stage="merozoite"
/lab_host="Bos taurus (cow)"
/clone_lib="Theileria annulata merozoite"
/notes="country: Turkey; Ankara"

ORIGIN

Oy		8964	AGATCTCAACAGGAATAATGGTGGCTAATGTCGAAATTGTGTGGAAATGTTGTAATGA	9023
Oy		8964		9023
Db		640	AAAGAAA	699
Oy		9024	TAAATGTATTAGAATAATGTAGAACCCAAATCGAGACATGCTATTGTAGAAACTAC	9083
Db		700	AAAAAAAAAGAAA	759
Oy		9084	TGGAATCAAGATTTCTCAGCATTTTACAAGCAAAGTAAATTTCTTGGAAGAGGA	9143
Db		760	AAAAAAAAAAAAAAAAAGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAA	819
Oy		9144	CGCTGCAGCTGCAGCTGCAATATGAAATGTACACTTCCATGATGATGATATTTAAAA	9203
Db		820	AGAAAAAAGAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAA	879
Oy		9204	TTTGGCAAGCAGTAGATCTCTCTCAATTATACCAAAAAATTCAAAAATAATATTAC	9263
Db		880	AAAGAAAAAAGAAAAAAG	939
Oy		9284	TTTAGCATCAAGTAGTGAATGCAATGTGTAATGTTATGCGGTGGCTGAAGAGGTGC	9323
Db		940	AAGAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA	999
Oy		9324	AGGAGCCAAAGCAGTGTAGTGTAAATCAATAATTAATGAACTAATTAATTTGATT	9383
Db		1000	AAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA	1059
Oy		9384	AGCAGGAAAAATTTAAACAGAGGAAACATCATGTATATGCCGATATGATAAAAATTA	9443
Db		1060	AAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	1119
Oy		9444	TAAATTAATGTAAGCAAAATTTTAAGGCTATTGCCGATCCAAAATCATCTGCGACTGC	9503
Db		1120	AAAGAAAAAAGAAAAA	1179
Oy		9504	TTCCGCAACTGCCACTATTGAAAAATTAATTAATGATCGATCCGAGATT	9563
Db		1180	AAAGAAAAAAGAAAAAAGAAAA	1239
Oy		9564	TAAAAATTAATCTGCGCAAGATTGGAAGGAAAGCTAATTAATAAAAGCTCGTAGACTTAA	9623
Db		1240	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	1299
Oy		9624	TCAGTAGACTGCTATACGATTAATATATACATGCGATTTCTTGAAAAAGCATCAAAA	9683
Db		1300	AAAAAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1359
Oy		9684	ATTGACATATCAATCAAGAGAGAGAAAAAGGAAAAA	9723
Db		1360	AGGAAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	1399
RESULT 39				
LOCUS	DRI41924	1783 bp	mRNA linear EST 16-JUN-2005	
DEFINITION	49190649 Drosophila pseudoobscura embryonic cDNA library Drosophila			
VERSION	pseudobsecura cDNA clone O7_3, mRNA sequence.			
KEYWORDS	DRI41924 .1 GI:67807014			
SOURCE	EST.			
ORGANISM	Drosophila pseudoobscura Drosophila pseudoobscura Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Empidoidea; Drosophilidae; Drosophila. 1 (bases 1 to 1783) Richardson, S., Liu, Y., Battencourt, B.R., Hradecky, P., Letovsky, S., Nielsen, R., Thornton, K., Hubisz, M.J., Chen, R., Meisel, R.P., van Batenburg, M.F., Howells, S.L., Scheier, S.E., Sodergren, E., Matthews, B.B., Crosby, M.A., Schroeder, A.J., Ortiz-Barrientos, D.,			

TITLE	JOURNAL	PUBMED	COMMENT
Rives,C.M., Weizker,M.L., Murny,D.M., Scott,G., Steffen,D., Wheeler,D.A., Worley,K.C., Havlik,P., Durbin,K.J., Egan,A., Gill,R., Hume,J., Morgan,M.B., Miner,G., Hamilton,C., Hang,Y., Waldron,L., Verdusco,D., Clerc-Blanchet,K.P., Dubochak,I., Noor,M.A., Anderson,W., White,K.P., Clark,A.G., Scheffer,S.W., Gelbart,W., Weinstein,G.M. and Gibbs,R.A.	Comparative genome sequencing of Drosophila pseudoobscura: Chromosomal, gene, and cis-element evolution	Genome Res. 15 (1), 1-18 (2005)	15632085
Contact: Stephen Richards Human Genome Sequencing Center Baylor College of Medicine One Baylor Plaza, Houston, TX 77030, USA Tel.: 713-798-6667 Email: stephen@bcm.tmc.edu NCBI Trace Archive: 226707030 Insert Length: 1750 Std Error: 0.25. Location/Qualifiers			
FEATURES			
Source			
1..1783	/organism="Drosophila pseudoobscura"		
	/mol_type="mRNA"		
	/db_xref="taxon:7237"		
	/clone="O7"		
	/dev_stage="0-18h embryos"		
	/clone_id="Drosophila pseudoobscura embryonic cDNA library"		
	/note="Vector: pOT87; Site 1: EcoRI; Site 2: XhoI; oligo dt priming from poly A+ RNA, directionally cloned"		
ORIGIN			
Query Match	0.98;	Score 87.2;	DB 8; Length 1783;
Best Local Similarity	43.1%;	Pred. No. 1.3e-06;	
Matches 468;	Conservative 0;	Mismatches 615;	Indels 3; Gaps 1;
Dy	8641 AAAAATGAATAATTATGGAATGTGAAAGAAGCACTCGAGACCCCTTAAGTCGCAGCCGAACA	8700	
Dd	402 AAAAAAAGAGAAAGGAAAAGGAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	461	
Dy	8701 GCAGCCGTTGAAAAATTATACAAAGAGTACTACAGACATTTGTTGACGAAATTTGGGAA	8760	
Dd	462 AAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	521	
Dy	8761 ATTGAGATTAATTAGAAACGATTGCAAGATTAATGCAATTGTAGAGTCAACGAGAC	8820	
Dd	522 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	581	
Dy	8821 GGAACCAAAGAGCTTGTTCGAAAGATGTGTTTTCTGTGAAAAATCATTTCAGCG	8880	
Dd	582 AAAAAAAAAAAGAAAAAAGAAAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	641	
Dy	8881 GAACCAAAATCATCCATTGAAGATPAAGCCAGATTTGTGGAACCGAAGTGAATGTA	8940	
Dd	642 AAAAAAAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	701	
Dy	8941 GATGCTTTGATGAACCTTAGTAGATCTACAGAAAAAGTGTGCTTAGTGTGCAATT	9000	
Dd	702 AAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	761	
Dy	9001 GGTTATTGAAATGTTGATGTGAATATGTGATTAGAAAAAATGTGAAGCCAAATTCGA	9060	
Dd	762 AAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	821	
Dy	9061 AGACATCTATTGTAGAACTACGAAAAAAGAAATATCAAGCATTTACAGAGCAAA	9120	
Dd	822 AAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	881	
Dy	9121 GTAATATTCTTGAAAAAGAGACGCTGCAGCTGCAGCTCAATATGCAATGTACACTT	9180	
Dd	882 AAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	941	
Dy	9181 TCCAATGATGATATTAAAAATTTGCCAAAGCAGTATGCATCTTCAATTAAATACC	9240	

Db	942	AAAAAAAAAAAAAAAAAAAAA---GAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	998
Qy	9241	AAAAATTCAAAAAATATATTAATCTTAGCATCAGTAGTGAATCGAATGTAATGTTTCAT	93000
Db	999	AAAGAAAAAGAAAAA	10588
Qy	9301	GGGGTGCCTGAAGCAAGAGTGTCAGACCCAAAGCAGATTAGTAAAGATCAATA	93606
Db	1059	AAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	11188
Qy	9361	AATAGAACTAATTAATGTTGATTTAGCAGAGAAAAATTAAACAGAGCGAAATCAATGTA	94208
Db	1119	AGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	11788
Qy	9421	TATGCCGATATGATAAATTTATATATATTAAGTAAGCAAAATCTAAGCGATTGGCAT	94808
Db	1179	AAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAGAG	12388
Qy	9481	GCCAAATCATGCTGCAGCTGCTTCCGCACTGCCATATTGAAGAAAAATGAAGTAAA	95408
Db	1239	AAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	12988
Qy	9541	TTTAATAATGCGATCCGAGAAATTTAAAAATATCTGCAAGATTGGAAGGAAAGCTAAT	96008
Db	1299	AAAAAGAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	13588
Qy	9601	AAAAAAACGTGCGTGAATCTTATCAGGTAGACTGTTACGATTAATATATACATGCGAT	96608
Db	1359	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAA	14188
Qy	9661	TCTTCTGAAAAAGCATACAAAAATTGACATATCATCAAGAGAGAGAAAAAGGAA	97208
Db	1419	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAA	14788
Qy	9721	AAATGA 9726	
Db	1479	AGAAAG 1484	

```

FEATURES
source
1..1650
location/Qualifiers
1..1650

LOCUS
CL041163

DEFINITION
Xenopus tropicalis (western clawed frog)

ACCESSION
CL041163

VERSION
CL041163.1

KEYWORDS
GSS.

SOURCE
Xenopus tropicalis (western clawed frog)

ORGANISM
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1650)
Kmitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 238
High quality sequence stop: 581.
location/Qualifiers
1..1650

```

```

/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-51P23"

```

```

/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_id="CH216"
/note="Vector: pTARAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

```

[illegible]

```

Db      1058 AAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAACAGAGAAAAACA 1117
QY      9381 TTTAGAGGAAAAATTAAAAACAGAGGAAA CATCAATGTATATGCCGATATGATAAAA 9440
Db      1118 AAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAACAAAAACAAAAACAAAAAG 1177
QY      9441 TTTATATATATAGTACAAATTCTAAGGCTATTGCCGATGCCAAAAGTCATGCTGCAGC 9500
Db      1178 AGAACAAAAAGAAAAAGAACGACAAACAAAAAGAAAAAGAAAAAGAAAAAG 1237
QY      9501 TGCTTCGCACTGCCATTTGAAAAAATGAAGTAAATTATTAATGCGATCGAGA 9560
Db      1238 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGAA 1297
QY      9561 ATTTAAAAATATCTGGCAAGATTGGAAGGAAAGCTAATAAAAAAACGTGCTAGGATC 9620
Db      1298 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAA 1357
QY      9621 TAATCAGGTAGCTGGTATCGGATTAATATACATGCGATTCTTGAAAAAGCATACAA 9680
Db      1358 AAAAGACAAAAACAAAAACAAAGAAAAAGAAAAAGAAAAAGAAAAATACCAAG 1417
QY      9681 AAAATTGACATATCAATCAAGAGAGAAAAAGGAAAAAA 9723
Db      1418 AAAACAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1460

```

Search completed: November 25, 2005, 20:31:17
 Job time : 22535 sec8

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2005, 14:17:17 ; Search time 4501 Seconds
(without alignments)
17868.911 Million cell updates/sec

Title: US-10-647-057-8

Perfect score: 9726
Sequence: 1 atgagcgccatcaataaataa.....gagaaaggaataatga 9726

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA.Main.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9726	100.0	9726	US-09-841-786-8	Sequence 8, Appl1
2	9726	100.0	9726	US-10-647-057-8	Sequence 8, Appl1
3	9726	100.0	11130	US-09-841-786-15	Sequence 15, Appl1
4	9726	100.0	11130	US-10-647-057-15	Sequence 15, Appl1
5	2780	28.6	2780	US-09-841-786-10	Sequence 10, Appl1
6	2780	28.6	2780	US-10-647-057-10	Sequence 10, Appl1
7	2322	23.9	2322	US-09-841-786-13	Sequence 13, Appl1
8	2322	23.9	2322	US-10-647-057-13	Sequence 13, Appl1
9	2141	22.0	2141	US-09-841-786-11	Sequence 11, Appl1
10	2141	22.0	2141	US-10-647-057-11	Sequence 11, Appl1
11	1887	19.4	1887	US-09-841-786-12	Sequence 12, Appl1
12	1887	19.4	1887	US-10-647-057-12	Sequence 12, Appl1
13	1130	11.6	1130	US-09-841-786-9	Sequence 9, Appl1
14	1130	11.6	1130	US-10-647-057-9	Sequence 9, Appl1
15	116.2	1.2	5152	US-09-742-096-2	Sequence 2, Appl1
16	116.2	1.2	5152	US-09-742-096-1	Sequence 1, Appl1
17	107.6	1.1	5528	US-10-415-253-1	Sequence 1, Appl1
18	80.4	0.8	1300	US-10-668-749A-1	Sequence 1, Appl1
19	79.2	0.8	7434	US-09-815-242-4761	Sequence 4761, Ap
20	79.2	0.8	7434	US-10-470-048B-201	Sequence 201, App
21	79.2	0.8	7437	US-09-815-242-8869	Sequence 8869, Ap
22	79.2	0.8	8155	US-08-781-986A-63	Sequence 63, Appl
23	79.2	0.8	8155	US-10-329-624-63	Sequence 63, Appl

24	74.8	0.8	7104	US-09-815-242-4580	Sequence 4580, Ap
25	74.8	0.8	7107	US-09-815-242-8291	Sequence 8291, Ap
26	72.2	0.8	600	US-10-972-079-13288	Sequence 13288, A
27	72.4	0.7	7446	US-10-282-122A-7578	Sequence 7578, Ap
28	70.6	0.7	11130	US-10-425-115-132399	Sequence 132399, Ap
29	68.8	0.7	31124	US-10-087-192-463	Sequence 463, App
30	68.2	0.7	4911	US-10-470-048B-427	Sequence 427, App
31	68.2	0.7	4997	US-10-282-122A-35506	Sequence 35506, A
32	67.6	0.7	7442	US-10-221-714A-409	Sequence 409, App
33	64.8	0.7	1891	US-09-742-096-4	Sequence 4, Appl1
34	64.4	0.7	1712	US-10-415-253-3	Sequence 3, Appl1
35	63.8	0.7	1121	US-10-425-115-115706	Sequence 115706, A
36	63.2	0.6	996	US-10-425-115-151933	Sequence 151933, A
37	62.6	0.6	673	US-10-424-599-57885	Sequence 57885, A
38	62.4	0.6	1062	US-10-425-115-120013	Sequence 120013, A
39	62.4	0.6	1243	US-10-705-635-24	Sequence 172717, A
40	62.4	0.6	50000	US-10-021-323-9336	Sequence 24, Appl
41	62.2	0.6	627	US-10-450-763-11257	Sequence 9336, Ap
42	61.6	0.6	441	US-10-311-455-1670	Sequence 11257, A
43	61.4	0.6	6668	US-10-311-455-2231	Sequence 1670, Ap
44	61.2	0.6	12237	US-10-311-455-2231	Sequence 2231, Ap
45	61	0.6	626	US-10-357-930-60960	Sequence 60960, A

ALIGNMENTS

```
RESULT 1
US-09-841-786-8
; Sequence 8, Application US/09841786
; Patent No. US20020054883A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 9726
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-8

Query Match      100.0%; Score 9726; DB 3; Length 9726;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGGCATCAAAATTAACGTTCAAGAGCAAGAAAGATACAGATTCTAAATAA 60
    |||
DB 1 ATGAGCGGCATCAAAATTAACGTTCAAGAGCAAGAAAGATACAGATTCTAAATAA 60
    |||

QY 61 GTTTTAAATGATTTGGGATTTGATTAACATATGACGTTGATGATGATGATGATGAT 120
    |||
DB 61 GTTTTAAATGATTTGGGATTTGATTAACATATGACGTTGATGATGATGATGATGAT 120
    |||

QY 121 ACCGCGACTGAGATTTTGAACAAATAATGAAAAAAGATATGTTTATGACTTACT 180
    |||
DB 121 ACCGCGACTGAGATTTTGAACAAATAATGAAAAAAGATATGTTTATGACTTACT 180
    |||

QY 181 ACAAACAATTTCAAGGAGAACGTTTAAACAGTTTAAATGATTTGCTTAAACAGAA 240
    |||
DB 181 ACAAACAATTTCAAGGAGAACGTTTAAACAGTTTAAATGATTTGCTTAAACAGAA 240
    |||

QY 241 AATAATATGACAAATATATATTTGGGAAAAAGATATGATCGGGGTAATATATCTTTT 300
    |||
DB 241 AATAATATGACAAATATATATTTGGGAAAAAGATATGATCGGGGTAATATATCTTTT 300
    |||
```

Db 241 AATAATATGCAAAATCTATATTTTGGGAAAAAGAAATAGTACGGGGTAATATCTTTT 300
Oy 301 AACTTGTCAATGAAAAATTTGAAGTAGATGGATTTATCAACGGAATTCGAGAAATATAA 360
Db 301 AACTTGTCAATGAAAAATTTGAAGTAGATGGATTTATCAACGGAATTCGAGAAATATAA 360
Oy 361 ATTGGAGGAATTTATTTCTTAAAGCTCGAAGGAGATGCGATGAGAAAAATGAGATT 420
Db 361 ATTGGAGGAATTTATTTCTTAAAGCTCGAAGGAGATGCGATGAGAAAAATGAGATT 420
Oy 421 ATCAATGCGGTTCTTTCAATCTATTTATCCAAAACAAGATATTTTAAAGAGCTTTG 480
Db 421 ATCAATGCGGTTCTTTCAATCTATTTATCCAAAACAAGATATTTTAAAGAGCTTTG 480
Oy 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGGAATCATCTCAGATGATGAAAAATGAAA 540
Db 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGGAATCATCTCAGATGATGAAAAATGAAA 540
Oy 541 ATTCCATTTGAATCCGAATGGAAGCATTAAGGTGAAGAAAAATCAATGCTGTTGAAGC 600
Db 541 ATTCCATTTGAATCCGAATGGAAGCATTAAGGTGAAGAAAAATCAATGCTGTTGAAGC 600
Oy 601 ATCGGTTTAAATGCGCGCGGATTTAGATTTGAAGATCATGCAATCTTAAGAAGCAAGATT 660
Db 601 ATCGGTTTAAATGCGCGCGGATTTAGATTTGAAGATCATGCAATCTTAAGAAGCAAGATT 660
Oy 661 ACAGATTTTAAATTTAGTCAATTAATGATCGAATTAATTTCTGCTGCGAGAT 720
Db 661 ACAGATTTTAAATTTAGTCAATTAATGATCGAATTAATTTCTGCTGCGAGAT 720
Oy 721 TTTAAAGCTACCAAGACAAATCTGAGATTAATTTCTTCACTCAGTCAATAGATTCTCT 780
Db 721 TTTAAAGCTACCAAGACAAATCTGAGATTAATTTCTTCACTCAGTCAATAGATTCTCT 780
Oy 781 CAAAAGCTATGCGGAAAAATTTCACTGTTGGAAGAAATGAAATATGTAAGAAAGCA 840
Db 781 CAAAAGCTATGCGGAAAAATTTCACTGTTGGAAGAAATGAAATATGTAAGAAAGCA 840
Oy 841 AATACCAAGCAAAATATTTGAATCTGATCTGTAATGGAAGCAGATGAAATATAAAATT 900
Db 841 AATACCAAGCAAAATATTTGAATCTGATCTGTAATGGAAGCAGATGAAATATAAAATT 900
Oy 901 AGTGGAAAGCTACAAATGAGATTTATAAAGAAAGGAGAAAAAACTTATTAAC 960
Db 901 AGTGGAAAGCTACAAATGAGATTTATAAAGAAAGGAGAAAAAACTTATTAAC 960
Oy 961 ACTCCTTAAAGTTTACAGATGATGAGATCTCCGTAAGATTAATTAAGAAAGTCAATA 1020
Db 961 ACTCCTTAAAGTTTACAGATGATGAGATCTCCGTAAGATTAATTAAGAAAGTCAATA 1020
Oy 1021 GGAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTATTAAT 1080
Db 1021 GGAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTATTAAT 1080
Oy 1081 AAGCTTGCAAAAGCACTTTTAACTTTGTTACAGGTTCTAATTTCTCATCAATTTAAT 1140
Db 1081 AAGCTTGCAAAAGCACTTTTAACTTTGTTACAGGTTCTAATTTCTCATCAATTTAAT 1140
Oy 1141 GGAATTTTAAAGTTTATGCAAGATGATGCAAGTCTGTTATTTGGAAGAAAGTCCAAATC 1200
Db 1141 GGAATTTTAAAGTTTATGCAAGATGATGCAAGTCTGTTATTTGGAAGAAAGTCCAAATC 1200
Oy 1201 GAAGCAACAGAAAGAAAGCAAAATTTCAATTTCTTACAGTGAAGTAAAGCAATGAGGA 1260
Db 1201 GAAGCAACAGAAAGAAAGCAAAATTTCAATTTCTTACAGTGAAGTAAAGCAATGAGGA 1260
Oy 1261 GCAGCTACTTCTCAATTAATAATTTCAATTTATTTGAGAAAGCCAAATGAGAAATCTT 1320
Db 1261 GCAGCTACTTCTCAATTAATAATTTCAATTTATTTGAGAAAGCCAAATGAGAAATCTT 1320
Oy 1321 CTCAGTATCGAGCGGAGATATTTCTGCAAAAGATTAATTCGATGTAATTTGAGAGA 1380
Db 1321 CTCAGTATCGAGCGGAGATATTTCTGCAAAAGATTAATTCGATGTAATTTGAGAGA 1380

Oy 1381 GAAGTAAAAATCGAAGGAGAGACAGATTTACTTCAAAATCTGAAAAATTAATGATGCT 1440
Db 1381 GAAGTAAAAATCGAAGGAGAGAGACAGATTTACTTCAAAATCTGAAAAATTAATGATGCT 1440
Oy 1441 TCTGTTCTGTTGGAACGATGAGAGATTTCAATTAAGTACTTTTCAGATTTGATGAG 1500
Db 1441 TCTGTTCTGTTGGAACGATGAGAGATTTCAATTAAGTACTTTTCAGATTTGATGAG 1500
Oy 1501 GAAGGAGAAATTAATCTTCGATCAAGATTCGTAAGAGCAAAATGATGATTCAGAAAACG 1560
Db 1501 GAAGGAGAAATTAATCTTCGATCAAGATTCGTAAGAGCAAAATGATGATTCAGAAAACG 1560
Oy 1561 GATGATGTAATTTGAGAGAGTGAAGCAATTAATTTCAATTCGAGCTGTTGAAAAAGTGA 1620
Db 1561 GATGATGTAATTTGAGAGAGTGAAGCAATTAATTTCAATTCGAGCTGTTGAAAAAGTGA 1620
Oy 1621 TTGGGGGATGATGATTTGAGGATTTGAGCAAAATTTCTAACTATATGATGCTTCTCC 1680
Db 1621 TTGGGGGATGATGATTTGAGGATTTGAGCAAAATTTCTAACTATATGATGCTTCTCC 1680
Oy 1681 CGTATGATGTAATTTGATATCTACATGCAAGAAAGCACTAAATTTGAGAGGCTCATTAAC 1740
Db 1681 CGTATGATGTAATTTGATATCTACATGCAAGAAAGCACTAAATTTGAGAGGCTCATTAAC 1740
Oy 1741 ATTACTAAAAATTAAGTCTTCTGCAAAACAGATCTGATTTGGGAATTTCCAGTTTATGAAT 1800
Db 1741 ATTACTAAAAATTAAGTCTTCTGCAAAACAGATCTGATTTGGGAATTTCCAGTTTATGAAT 1800
Oy 1801 GATCACTTTATGATATGATCATCTAATCAATTTTATGATGCAATTAACACGCGGTTT 1860
Db 1801 GATCACTTTATGATATGATCATCTAATCAATTTTATGATGCAATTAACACGCGGTTT 1860
Oy 1861 GGAAGAGACAGTGTCAATGAGAAATTAAGATTAAGCACTAATTTATGATGATGCT 1920
Db 1861 GGAAGAGACAGTGTCAATGAGAAATTAAGATTAAGCACTAATTTATGATGATGCT 1920
Oy 1921 GTGCTGCAACCAATGAGCAAAATCATATTAATTTCTGCTCTGAGCAATGAGAGAGTGA 1980
Db 1921 GTGCTGCAACCAATGAGCAAAATCATATTAATTTCTGCTCTGAGCAATGAGAGAGTGA 1980
Oy 1981 AGACTTTCTTCAAGAGTGAAGGAGATGTAATGTAAGGCAATTAATTAAGGCTCAAAATCTT 2040
Db 1981 AGACTTTCTTCAAGAGTGAAGGAGATGTAATGTAAGGCAATTAATTAAGGCTCAAAATCTT 2040
Oy 2041 CGAGCGACTACGTAAGTGAAGTGTGCTGTACGAAAGGAAAGAAAGAAAGAACTTAAT 2100
Db 2041 CGAGCGACTACGTAAGTGAAGTGTGCTGTACGAAAGGAAAGAAAGAAAGAAAGCTTAAT 2100
Oy 2101 GGAATGCAAGATTTTATGAGAACTATTAATTAATGCTTCTGAGCAATTTGCCGAT 2160
Db 2101 GGAATGCAAGATTTTATGAGAACTATTAATTAATGCTTCTGAGCAATTTGCCGAT 2160
Oy 2161 CATGCTGAATGTTGATGGAAGGAAATTTGATCAACAGTGAAGAAATTAATTAATGATAT 2220
Db 2161 CATGCTGAATGTTGATGGAAGGAAATTTGATCAACAGTGAAGAAATTAATTAATGATAT 2220
Oy 2221 AAAAATCTTCAAAATGAGCAAAAGTCTGTTATTAATTAATTAAGAAAGAGCT 2280
Db 2221 AAAAATCTTCAAAATGAGCAAAAGTCTGTTATTAATTAATTAAGAAAGAGCT 2280
Oy 2281 TTTGAAAAAGAAAGAAAGCTCCAGAAATTAATGATCCGAAAGTATGTAATCTTAATGAAAA 2340
Db 2281 TTTGAAAAAGAAAGAAAGCTCCAGAAATTAATGATCCGAAAGTATGTAATCTTAATGAAAA 2340
Oy 2341 TTATTTGAATGATTTTCAAGAAATTTGATGAGAAACCGAGCTTTTACTAAATGATGAA 2400
Db 2341 TTATTTGAATGATTTTCAAGAAATTTGATGAGAAACCGAGCTTTTACTAAATGATGAA 2400
Oy 2401 AGAATGCAATTAATTTCTTCCGAGTGAAGCTTCAAAACAGGAATCTGCTATTAAGAAATGCA 2460
Db 2401 AGAATGCAATTAATTTCTTCCGAGTGAAGCTTCAAAACAGGAATCTGCTATTAAGAAATGCA 2460

Oy	2461	AACTATGCTTCAGGAGAAATGAAAAATTAAGAGAAAAATTCAGAAAGATTTAAAGCT	2520
Db	2461	AACTATGCTTCAGGAGAAATGAAAAATTAAGAGAAAAATTCAGAAAGATTTAAAGCT	2520
Oy	2521	TTTTCAAAAGATTGAGTGGCATGTATTAAAGAACTTTGAATTTTACAGAGTNGAAAT	2580
Db	2521	TTTTCAAAAGATTGAGTGGCATGTATTAAAGAACTTTGAATTTTACAGAGTNGAAAT	2580
Oy	2581	TATGCAAAATTTCAACTTTTACCTCTTCGAGACTAATGAGAAAGATGTTTCTCT	2640
Db	2581	TATGCAAAATTTCAACTTTTACCTCTTCGAGACTAATGAGAAAGATGTTTCTCT	2640
Oy	2641	GTGGAGAGAGCTGTTTCGTGGTNGAACAGAGAAATTATAGCAAGATATCCGTTGGAAAA	2700
Db	2641	GTGGAGAGAGCTGTTTCGTGGTNGAACAGAGAAATTATAGCAAGATATCCGTTGGAAAA	2700
Oy	2701	GGAGCTAAACTCTGCTGCAAAAAAGATTTTAAATTAAAGCTATCAATTAAGCAAGAAC	2760
Db	2701	GGAGCTAAACTCTGCTGCAAAAAAGATTTTAAATTAAAGCTATCAATTAAGCAAGAAC	2760
Oy	2761	GTGAATTTAGTTGGAAATATTGGACTTGCGAAGACGTACATCCGGAGTGCAGTCCGA	2820
Db	2761	GTGAATTTAGTTGGAAATATTGGACTTGCGAAGACGTACATCCGGAGTGCAGTCCGA	2820
Oy	2821	GGAAAGATTAAATGTTCCAAAGATCGAAAAATTCAGCTATCGTAGAGCTAAAGAAAAAGCT	2880
Db	2821	GGAAAGATTAAATGTTCCAAAGATCGAAAAATTCAGCTATCGTAGAGCTAAAGAAAAAGCT	2880
Oy	2881	GAAATTAACAGAGAAAAATTAATATGCAAGTCAATTTGAAAGCTATTAATTAAGCCGGA	2940
Db	2881	GAAATTAACAGAGAAAAATTAATATGCAAGTCAATTTGAAAGCTATTAATTAAGCCGGA	2940
Oy	2941	TCTTTAATGATGGCTCAGGTGGGGAAGCAATCAATGGAATGGGAGTTTATGAGAGGT	3000
Db	2941	TCTTTAATGATGGCTCAGGTGGGGAAGCAATCAATGGAATGGGAGTTTATGAGAGGT	3000
Oy	3001	ATCAGTAAGGCAGAGATTTCCATTGATGACGAGCAATTTGAAAGCTAATTAAAAAAATTT	3060
Db	3001	ATCAGTAAGGCAGAGATTTCCATTGATGACGAGCAATTTGAAAGCTAATTAAAAAAATTT	3060
Oy	3061	GCTTTAAACAGTAAGAAATGATCTTCTGTTTGGAAATGTCGCGGTTCAGCGCGGAATCCGA	3120
Db	3061	GCTTTAAACAGTAAGAAATGATCTTCTGTTTGGAAATGTCGCGGTTCAGCGCGGAATCCGA	3120
Oy	3121	ACGAAAAATGCGGCGGTGCGGGGTGCTGTGCGGTAATGATTAATTTCAAAACAA	3180
Db	3121	ACGAAAAATGCGGCGGTGCGGGGTGCTGTGCGGTAATGATTAATTTCAAAACAA	3180
Oy	3181	GCTTCATTGAAGATTAATGACGAGGACAAAGTAATATGATTAAGAAATTAAGATGATGAA	3240
Db	3181	GCTTCATTGAAGATTAATGACGAGGACAAAGTAATATGATTAAGAAATTAAGATGATGAA	3240
Oy	3241	GTAACAGTAACTGCGGAAATCTTTAGAGTAGATGCAAAAACGACCGGAACATCAACAGT	3300
Db	3241	GTAACAGTAACTGCGGAAATCTTTAGAGTAGATGCAAAAACGACCGGAACATCAACAGT	3300
Oy	3301	ATTTCTGTGTCGCGGAGGAATTAATTAAGTTGGAAGTAAACGAGTGAAGAAAAACGAAA	3360
Db	3301	ATTTCTGTGTCGCGGAGGAATTAATTAAGTTGGAAGTAAACGAGTGAAGAAAAACGAAA	3360
Oy	3361	TCAGAAAGAAAGACGAGAGGATTTTGTGGCAAAATCCGAAACAAAGTGAATCTGTAAAA	3420
Db	3361	TCAGAAAGAAAGACGAGAGGATTTTGTGGCAAAATCCGAAACAAAGTGAATCTGTAAAA	3420
Oy	3421	AATTAATAATTCGAGATATGATGATTCATTTAACGAAAAAATTCAAATTAATTTCTGAA	3480
Db	3421	AATTAATAATTCGAGATATGATGATTCATTTAACGAAAAAATTCAAATTAATTTCTGAA	3480
Oy	3481	GGAGTAAAAAAGCGGGGAATTTCTCTTCGAAAGCTTTCTCAATCTCCCGATTAAGAGACG	3540
Db	3481	GGAGTAAAAAAGCGGGGAATTTCTCTTCGAAAGCTTTCTCAATCTCCCGATTAAGAGACG	3540
Oy	3541	TCCTTCAGTTGGAGCTTCGGAAGTCTTCTTCAATAATATTTAAAAAGAAACAATCT	3600

Db	3541	TCCTTCAGTTGGAGGCTTCGGAAGTGTCTTCAATATATTTAAAGAAACACTCT	3600
QY	3601	GCTGTCGTAGATGAGATAAGATTAATTTGAAAGGAGCAATAAAGATAGAGTACT	3660
Db	3601	GCTGTCGTAGATGAGATTAAGATTAATTTGAAAGGAGCAATAAAGATAGAGTACT	3660
QY	3661	TCCTTCGATTCCTCTTTGTTGGAGCATGCGGCGGATCTGCTGCACTTCACTGGAATCAT	3720
Db	3661	TCCTTCGATTCCTCTTTGTTGGAGCATGCGGCGGATCTGCTGCACTTCACTGGAATCAT	3720
QY	3721	ATTGGAAGTGAATATACCAACATAGTGTGCTTTACCTGAGCGGCTGCTGAATAT	3780
Db	3721	ATTGGAAGTGAATATACCAACATAGTGTGCTTTACCTGAGCGGCTGCTGAATAT	3780
QY	3781	ATTCAAAAGTAAACAAGTCTTTGGTTAAAAATAGTATATTCCAAATGCCAAATAAATTT	3840
Db	3781	ATTCAAAAGTAAACAAGTCTTTGGTTAAAAATAGTATATTCCAAATGCCAAATAAATTT	3840
QY	3841	AAAGTAAATGCTTTGAGTGAAGAACTCAAGTACACAGAGAGCAAGTTTGGAGCACTT	3900
Db	3841	AAAGTAAATGCTTTGAGTGAAGAACTCAAGTACACAGAGAGCAAGTTTGGAGCACTT	3900
QY	3901	AAAGAAATGAGAGACAAAGAAAAGTTATCTAATGGAACTTCTGCTTCATACCTTA	3960
Db	3901	AAAGAAATGAGAGACAAAGAAAAGTTATCTAATGGAACTTCTGCTTCATACCTTA	3960
QY	3961	GTGAACAATGAAGTTCTGCAAAATCAAGAAATATACAGTACAGAGAGAACTGAAGC	4020
Db	3961	GTGAACAATGAAGTTCTGCAAAATCAAGAAATATACAGTACAGAGAGAACTGAAGC	4020
QY	4021	CAAAAAATGATGTTGATGTCATGCTTATCAAGCGACACCCAAAGTACAGAGCTTTA	4080
Db	4021	CAAAAAATGATGTTGATGTCATGCTTATCAAGCGACACCCAAAGTACAGAGCTTTA	4080
QY	4081	AATTTTACAGCTGGAAGTCAATGGAATCTAGGGGCTACTGTGACTGTTGCCAAATTA	4140
Db	4081	AATTTTACAGCTGGAAGTCAATGGAATCTAGGGGCTACTGTGACTGTTGCCAAATTA	4140
QY	4141	AACAACAAGTAAATGCTTCAATTAGTGTGAGGTGAGATATACTAATCGTTAATCGACCGAC	4200
Db	4141	AACAACAAGTAAATGCTTCAATTAGTGTGAGGTGAGATATACTAATCGTTAATCGACCGAC	4200
QY	4201	GCAAAAGCTCTTTAGCAACCACTCAAGTGACTGCTGCACTGACGACGCGGAGCAATT	4260
Db	4201	GCAAAAGCTCTTTAGCAACCACTCAAGTGACTGCTGCACTGACGACGCGGAGCAATT	4260
QY	4261	AGTTCTGAGCGGGATTGGAATTTATCAAGGGGCTGTTCTGTCAATAGAATTGACAT	4320
Db	4261	AGTTCTGAGCGGGATTGGAATTTATCAAGGGGCTGTTCTGTCAATAGAATTGACAT	4320
QY	4321	GACGTGGAAGCTAGGCTGATTAATCTTCCATCGAAGAGCTTAATGAATTAATGTCAAT	4380
Db	4321	GACGTGGAAGCTAGGCTGATTAATCTTCCATCGAAGAGCTTAATGAATTAATGTCAAT	4380
QY	4381	GCCAAGAATGTCAAAGGAAGTCTGATCTACCAAAAGATATCAGGCTTTACTAAATGGA	4440
Db	4381	GCCAAGAATGTCAAAGGAAGTCTGATCTACCAAAAGATATCAGGCTTTACTAAATGGA	4440
QY	4441	AAAAGTAAAAATATTTGAAAGATCGTGTATTAAATGACCTGAAATAGGTTAATATACG	4500
Db	4441	AAAAGTAAAAATATTTGAAAGATCGTGTATTAAATGACCTGAAATAGGTTAATATACG	4500
QY	4501	AAAGAAACAATGAGAAAAGCAAGAAAAGAAAAGAGAGCGCTCATGTGAAAATGCTCTTTA	4560
Db	4501	AAAGAAACAATGAGAAAAGCAAGAAAAGAAAAGAGAGCGCTCATGTGAAAATGCTCTTTA	4560
QY	4561	TCGGTTGCTGGAACGATTAATCCGCTGGAAGATAGCTAATTCGACGTAATACGTTAAA	4620
Db	4561	TCGGTTGCTGGAACGATTAATCCGCTGGAAGATAGCTAATTCGACGTAATACGTTAAA	4620
QY	4621	AATTAATTTAAAGCAATGAGTGAAGCAATAAGAAAGCCGAGAGGATTAATTCAT	4680
Db	4621	AATTAATTTAAAGCAATGAGTGAAGCAATAAGAAAGCCGAGAGGATTAATTCAT	4680

Db 4621 AATAAATTTAAAGCAATTGAGTGAGAACCAATTAAGAAACCGAGAGATTAATTCAT 4680
Qy 4681 GCGAACAATGTAATGTGAGAGCAAAATCATCTACTGTGTGTGAATGCGGCTTCTGGA 4740
Db 4681 GCGAACAATGTAATGTGAGAGCAAAATCATCTACTGTGTGTGAATGCGGCTTCTGGA 4740
Qy 4741 CTTCGTATCAGCAAAAGATGCTTTTTCAGGAATGGATCTGAGCATGCGAAGACTTTATCA 4800
Db 4741 CTTCGTATCAGCAAAAGATGCTTTTTCAGGAATGGATCTGAGCATGCGAAGACTTTATCA 4800
Qy 4801 AATGACACGATTGCAAAAGTGATAAGGAAGAAATTTCTGCTGATTCCTTTAATGTGAAC 4860
Db 4801 AATGACACGATTGCAAAAGTGATAAGGAAGAAATTTCTGCTGATTCCTTTAATGTGAAC 4860
Qy 4861 GCAAAATTAATTCATTTCTGGGGTGAATGTTCGGGGAACCATGCGGTTCTCTTCTACG 4920
Db 4861 GCAAAATTAATTCATTTCTGGGGTGAATGTTCGGGGAACCATGCGGTTCTCTTCTACG 4920
Qy 4921 GCGGTGAGAGCTGCTTTTTCGAAATATACCTTCATATATAAACCCTGCTTGTATACA 4980
Db 4921 GCGGTGAGAGCTGCTTTTTCGAAATATACCTTCATATATAAACCCTGCTTGTATACA 4980
Qy 4981 GGAAGCAAGAGTAAATCCCTTTTGTGGAAGAAATACAAAGTCAATGTAACGCTTGAAT 5040
Db 4981 GGAAGCAAGAGTAAATCCCTTTTGTGGAAGAAATACAAAGTCAATGTAACGCTTGAAT 5040
Qy 5041 GATTCATATATTACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAGCAGGCTGGAATC 5100
Db 5041 GATTCATATATTACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAGCAGGCTGGAATC 5100
Qy 5101 GGAGGAATGTAATCTGTCAATCGTGTCTGTATGAACGGAAGCTTTAGTTAGTATCT 5160
Db 5101 GGAGGAATGTAATCTGTCAATCGTGTCTGTATGAACGGAAGCTTTAGTTAGTATCT 5160
Qy 5161 GAGTTGAAGAGTAAGTCTTTCATATGTAATGCAAAAGTCAAAAGCAATTAATACA 5220
Db 5161 GAGTTGAAGAGTAAGTCTTTCATATGTAATGCAAAAGTCAAAAGCAATTAATACA 5220
Qy 5221 ATTGCGCAAAATGCAAAATGAGAAAGCGCTGAGTGGAGCAACAGTTGCTATACA 5280
Db 5221 ATTGCGCAAAATGCAAAATGAGAAAGCGCTGAGTGGAGCAACAGTTGCTATACA 5280
Qy 5281 AATATTGGAAAAATCACTTAATAGTATTGTAAAAACAGTAAATTAACAACGCGAAT 5340
Db 5281 AATATTGGAAAAATCACTTAATAGTATTGTAAAAACAGTAAATTAACAACGCGAAT 5340
Qy 5341 GATCAAGATGAAAAAATATCAATGTGACTGCAAAAAGTATATATGACAAATCTATA 5400
Db 5341 GATCAAGATGAAAAAATATCAATGTGACTGCAAAAAGTATATATGACAAATCTATA 5400
Qy 5401 GCAGTCGAGTGGAGAGCAAAAGAGCCTCTGTGCAAGAGACTTCTGCAAGTACTACC 5460
Db 5401 GCAGTCGAGTGGAGAGCAAAAGAGCCTCTGTGCAAGAGACTTCTGCAAGTACTACC 5460
Qy 5461 TTGAATTAAGACAGTTTCTTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGAA 5520
Db 5461 TTGAATTAAGACAGTTTCTTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGAA 5520
Qy 5521 GAAATTAATGCAAAATGAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATCGAGTCA 5580
Db 5521 GAAATTAATGCAAAATGAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATCGAGTCA 5580
Qy 5581 GTGGTCACAAAATGCGACAGTGTCTTCGAGCAAGTGGACAAAGCTGCACTAGAGACTGGA 5640
Db 5581 GTGGTCACAAAATGCGACAGTGTCTTCGAGCAAGTGGACAAAGCTGCACTAGAGACTGGA 5640
Qy 5641 GTAGCAGTTAATTAATTAACAAAAATCTTCTGCAATATATAAATATGTAAGTCAAAAT 5700
Db 5641 GTAGCAGTTAATTAATTAACAAAAATCTTCTGCAATATATAAATATGTAAGTCAAAAT 5700
Qy 5701 GTACGAAATGCTTTGTGTAAAAAGCAAAATCTCATTTATTAACAAATTTGAAATTTGGA 5760
Db 5701 GTACGAAATGCTTTGTGTAAAAAGCAAAATCTCATTTATTAACAAATTTGAAATTTGGA 5760

Qy 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGATGACAGGTTCTGTAGCAGTGAATTAAGATTGTA 5820
Db 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGATGACAGGTTCTGTAGCAGTGAATTAAGATTGTA 5820
Qy 5821 AATAATACGATGACGAATTAATCATGCAAAAATCACTGCGAAAGGAAATGTGGAATT 5880
Db 5821 AATAATACGATGACGAATTAATCATGCAAAAATCACTGCGAAAGGAAATGTGGAATT 5880
Qy 5881 ATTACAGAGCTGATGCGGTAAATGCTAATTAATGACAGAAACAGTGTCTGAGTGGCCGT 5940
Db 5881 ATTACAGAGCTGATGCGGTAAATGCTAATTAATGACAGAAACAGTGTCTGAGTGGCCGT 5940
Qy 5941 GCACCAATAGAGGCTCAACCAAGTGTGAATGAATTAACAGATCTACAAAGCATATGTA 6000
Db 5941 GCACCAATAGAGGCTCAACCAAGTGTGAATGAATTAACAGATCTACAAAGCATATGTA 6000
Qy 6001 AAAATCTTACAGTATTGCTTAAGAAAGAAACAGATGTTATTTACTACTCAAGGCGAA 6060
Db 6001 AAAATCTTACAGTATTGCTTAAGAAAGAAACAGATGTTATTTACTACTCAAGGCGAA 6060
Qy 6061 GTAGTAAAGTGATGATTAAGTAAATGTAATCAAAATCTTAATTAACGAAGACTTATCAAA 6120
Db 6061 GTAGTAAAGTGATGATTAAGTAAATGTAATCAAAATCTTAATTAACGAAGACTTATCAAA 6120
Qy 6121 AAAAGAAAAATTAATTAATTAAGAAAGATTTGTTACCAATAGTTCAGCTACTCATCTTTA 6180
Db 6121 AAAAGAAAAATTAATTAATTAAGAAAGATTTGTTACCAATAGTTCAGCTACTCATCTTTA 6180
Qy 6181 AAATCTTTATTGGCAATTCGCGTGTTCAGAGCAACAGCGGAGTGGCAGAACTGTTAAT 6240
Db 6181 AAATCTTTATTGGCAATTCGCGTGTTCAGAGCAACAGCGGAGTGGCAGAACTGTTAAT 6240
Qy 6241 ATCAACAAAGTTTATGAGAAACAGAAAGCTCTGTGAGAAATTTCTATTAATGCAAAA 6300
Db 6241 ATCAACAAAGTTTATGAGAAACAGAAAGCTCTGTGAGAAATTTCTATTAATGCAAAA 6300
Qy 6301 CATTATCTGTAAATCAGAGATTAACAGAAATTCATTCGAGATAGTATGTTCTGTGGT 6360
Db 6301 CATTATCTGTAAATCAGAGATTAACAGAAATTCATTCGAGATAGTATGTTCTGTGGT 6360
Qy 6361 GTTGTGGAATGTAGAGATGAGAGCTTCTCTGTATACCAATTTATTAAGAAATACC 6420
Db 6361 GTTGTGGAATGTAGAGATGAGAGCTTCTCTGTATACCAATTTATTAAGAAATACC 6420
Qy 6421 AAGCAAGAGTTGAAAAAATCAATGTCGATGTAAGGTTTCGGAGAAAGCTGAAATT 6480
Db 6421 AAGCAAGAGTTGAAAAAATCAATGTCGATGTAAGGTTTCGGAGAAAGCTGAAATT 6480
Qy 6481 ACAGCAGATTCTAAGCAAGAAATTTCTTTTGAAGTGGAGTCCGACAGCCGGGTA 6540
Db 6481 ACAGCAGATTCTAAGCAAGAAATTTCTTTTGAAGTGGAGTCCGACAGCCGGGTA 6540
Qy 6541 GAGCGCGAGTGGCAGGAACCGTTCCGTAAATCAATTTGCGAGGAAGAGAGTAAAT 6600
Db 6541 GAGCGCGAGTGGCAGGAACCGTTCCGTAAATCAATTTGCGAGGAAGAGAGTAAAT 6600
Qy 6601 GTGGAAGAGCAAAAGTTTGTATAAABAACTGAGATTACAGCAAAAAGTTATAGTCT 6660
Db 6601 GTGGAAGAGCAAAAGTTTGTATAAABAACTGAGATTACAGCAAAAAGTTATAGTCT 6660
Qy 6661 GTTGCAATTTGAAAAATGCGCAGTGGAGTGGCTCAAAAGAGCTGGAATTTGAGACACA 6720
Db 6661 GTTGCAATTTGAAAAATGCGCAGTGGAGTGGCTCAAAAGAGCTGGAATTTGAGACACA 6720
Qy 6721 GTGGCAGTTACCAAGATGAATCAACACGAGACAGAGTGAATAATTTCTAAAATTATG 6780
Db 6721 GTGGCAGTTACCAAGATGAATCAACACGAGACAGAGTGAATAATTTCTAAAATTATG 6780
Qy 6781 ACTGGAACCAAGTTAGATGTAATAGCAGAAATGAGATTAATTAACAGGATCTGGAATGGGT 6840
Db 6781 ACTGGAACCAAGTTAGATGTAATAGCAGAAATGAGATTAATTAACAGGATCTGGAATGGGT 6840

QY 6841 TCAGCCGAGCTGGAATTTCTTGACGCCGAGATCTGAGAGCTTTCTGCAATAATATT 6900
 DB 6841 TCAGCCGAGCTGGAATTTCTTGACGCCGAGATCTGAGAGCTTTCTGCAATAATATT 6900
 QY 6901 GCAAAATTAAGTGAAGACAGATATCGATCATAGTACTTTACTTCTACTGATTAAT 6960
 DB 6901 GCAAAATTAAGTGAAGACAGATATCGATCATAGTACTTTACTTCTACTGATTAAT 6960
 QY 6961 GTAAAGCTCTTAATTAATTTGAAATTCCTTGACAGCCGGTGAAGAGCCGAGCTCTT 7020
 DB 6961 GTAAAGCTCTTAATTAATTTGAAATTCCTTGACAGCCGGTGAAGAGCCGAGCTCTT 7020
 QY 7021 GCAGAGTTACCGGAGTGTCTTCTGTAACATTAATTAATGTTCTGTGATAGCTGAGT 7080
 DB 7021 GCAGAGTTACCGGAGTGTCTTCTGTAACATTAATTAATGTTCTGTGATAGCTGAGT 7080
 QY 7081 CACAATTAATCTGATTTGACTCCGGTACGAGAAAAGTAAATGTAACGGCAAAAGAGNA 7140
 DB 7081 CACAATTAATCTGATTTGACTCCGGTACGAGAAAAGTAAATGTAACGGCAAAAGAGNA 7140
 QY 7141 AAAAAATATTAAAGCAACAGCAGCAATGCGAGAAATCGAGAGCAGCAATCGAGCCAA 7200
 DB 7141 AAAAAATATTAAAGCAACAGCAGCAATGCGAGAAATCGAGAGCAGCAATCGAGCCAA 7200
 QY 7201 GTCTTGTTAATTAATTTTGGAAACAGCTGTGAAGATGAAGAAATTTCTGAAGAAAGGA 7260
 DB 7201 GTCTTGTTAATTAATTTTGGAAACAGCTGTGAAGATGAAGAAATTTCTGAAGAAAGGA 7260
 QY 7261 ACAGAGTTTTTAAACCTTTAGACGAAGTTAAAGAAACAGATTAAGAAATTAATGAT 7320
 DB 7261 ACAGAGTTTTTAAACCTTTAGACGAAGTTAAAGAAACAGATTAAGAAATTAATGAT 7320
 QY 7321 GCTACGAAAAAATCTTACAAATCAGCAGGTAATTTCTACAGAAATCTTCTGTAAAGCG 7380
 DB 7321 GCTACGAAAAAATCTTACAAATCAGCAGGTAATTTCTACAGAAATCTTCTGTAAAGCG 7380
 QY 7381 GATTAAGAGATCTCAGGAGAGAGAAATTAAGCCATTGTGAAGACTTCTGATTTATT 7440
 DB 7381 GATTAAGAGATCTCAGGAGAGAGAAATTAAGCCATTGTGAAGACTTCTGATTTATT 7440
 QY 7441 GGAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7500
 DB 7441 GGAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7500
 QY 7501 GGAAGATGAGCTCTTCTGCTCCGATCAGGAAACAGTGGCAGTTAAACAATTTAAAGAAAT 7560
 DB 7501 GGAAGATGAGCTCTTCTGCTCCGATCAGGAAACAGTGGCAGTTAAACAATTTAAAGAAAT 7560
 QY 7561 TCCGAGATTACTGTTGAAAAATCTTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGATCG 7620
 DB 7561 TCCGAGATTACTGTTGAAAAATCTTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGATCG 7620
 QY 7621 GATTAATTAAGGAAATGTTCTTAAACAGATTAATCAAGGTCCTGTAGAGCAATTTGGGAA 7680
 DB 7621 GATTAATTAAGGAAATGTTCTTAAACAGATTAATCAAGGTCCTGTAGAGCAATTTGGGAA 7680
 QY 7681 GAGAGCTGCTATGACAGAAATTAATTTCTAAATGGAAGATCAATATCATGATTTAAATTTCT 7740
 DB 7681 GAGAGCTGCTATGACAGAAATTAATTTCTAAATGGAAGATCAATATCATGATTTAAATTTCT 7740
 QY 7741 AAGCTATTAGAAAAAATATGATGTTATTTAAAGATTAATTCGAAATTTGAGAGCGGA 7800
 DB 7741 AAGCTATTAGAAAAAATATGATGTTATTTAAAGATTAATTCGAAATTTGAGAGCGGA 7800
 QY 7801 GCAAAAGATTAACGTTAGAGCGGTAGCTGCGGAGCCATTATCTCAAAAGCAAGAAAT 7860
 DB 7801 GCAAAAGATTAACGTTAGAGCGGTAGCTGCGGAGCCATTATCTCAAAAGCAAGAAAT 7860
 QY 7861 GAAATGAATTCAGAGGTTGAAATTTGAGAGAGATATTTTCAATGAGAAAAATAGATTAAT 7920
 DB 7861 GAAATGAATTCAGAGGTTGAAATTTGAGAGAGATATTTTCAATGAGAAAAATAGATTAAT 7920
 QY 7921 AGCCCTTTAAAGAAATTTGAGAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGTC 7980

DB 7921 AGCCCTTTAAAGAAATTTGAGAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGTC 7980
 QY 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTACGAGGGCAGAAATTTCCGAA 8040
 DB 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTACGAGGGCAGAAATTTCCGAA 8040
 QY 8041 GCAAAAGATGCGGAAAGCTCTTAATTTGAAAGTTAGTCAAAATCCGGAAGAAATTTT 8100
 DB 8041 GCAAAAGATGCGGAAAGCTCTTAATTTGAAAGTTAGTCAAAATCCGGAAGAAATTTT 8100
 QY 8101 CATGAGATTAATGTAATTAAGAGCAACACATAAAATGAAGTAAACAGAGTTCTAAA 8160
 DB 8101 CATGAGATTAATGTAATTAAGAGCAACACATAAAATGAAGTAAACAGAGTTCTAAA 8160
 QY 8161 GCAGTAACAGGTTCTGATTTGGAGAGAGTGAAGTACCAAGCAGAACTGCTGCA 8220
 DB 8161 GCAGTAACAGGTTCTGATTTGGAGAGAGTGAAGTGAAGTCAAGAGCAATGCTGCA 8220
 QY 8221 GGTAAATCTATGTAAGAGTTGAGAGAGAAATTTGTTCAAGAACAAATCGAATGCA 8280
 DB 8221 GGTAAATCTATGTAAGAGTTGAGAGAGAAATTTGTTCAAGAACAAATCGAATGCA 8280
 QY 8281 ATTTCTAAAGTGAAGGTTTGAAGTGAAGTAAAGTAACTGCTAAATCTTCTGTAGTACA 8340
 DB 8281 ATTTCTAAAGTGAAGGTTTGAAGTGAAGTAAAGTAACTGCTAAATCTTCTGTAGTACA 8340
 QY 8341 GGAATGAGAGAGAAATTTCCGAGCAGAGTGAATTAATCTTACAGCAAAAGTAATCT 8400
 DB 8341 GGAATGAGAGAGAAATTTCCGAGCAGAGTGAATTAATCTTACAGCAAAAGTAATCT 8400
 QY 8401 GAATCCGTAATGCTTTACGAAAGCAAGATTAAGAAATTAATTAATTAATTAATTAAT 8460
 DB 8401 GAATCCGTAATGCTTTACGAAAGCAAGATTAAGAAATTAATTAATTAATTAATTAAT 8460
 QY 8461 GAATCCGTAATGCTTTACGAAAGCAAGATTAAGAAATTAATTAATTAATTAATTAAT 8520
 DB 8461 GAATCCGTAATGCTTTACGAAAGCAAGATTAAGAAATTAATTAATTAATTAATTAAT 8520
 QY 8521 TTAGCGGTAGCCGCTGTCATGCAACAGAAACAAAGCAATTTAGAGATCAAAACAG 8580
 DB 8521 TTAGCGGTAGCCGCTGTCATGCAACAGAAACAAAGCAATTTAGAGATCAAAACAG 8580
 QY 8581 TTAACCTTACAACTGTAAATGAGAGAAACGATCTCAACTTCGCAAAAGCTTGGCT 8640
 DB 8581 TTAACCTTACAACTGTAAATGAGAGAAACGATCTCAACTTCGCAAAAGCTTGGCT 8640
 QY 8641 AAAAAATGAATTAAGAAATGTAAGAACTGAGAGAGCTTGTAGTGGAGAGAAACA 8700
 DB 8641 AAAAAATGAATTAAGAAATGTAAGAACTGAGAGAGCTTGTAGTGGAGAGAAACA 8700
 QY 8701 GCAGCCGTTGAAGAAATTAACAAAGATCTACAGAGCATTTGTTGAGAAATTTGGGA 8760
 DB 8701 GCAGCCGTTGAAGAAATTAACAAAGATCTACAGAGCATTTGTTGAGAAATTTGGGA 8760
 QY 8761 ATTGAAGTAAATTTGAAACGATTCGACAGATTAATGATGTTAAGATCAACGAGAC 8820
 DB 8761 ATTGAAGTAAATTTGAAACGATTCGACAGATTAATGATGTTAAGATCAACGAGAC 8820
 QY 8821 GGAACCAAGAGGCTTGTGCGAAAGATGTTCTGTGTAAGAAATTAATTTCAATTTCAAGG 8880
 DB 8821 GGAACCAAGAGGCTTGTGCGAAAGATGTTCTGTGTAAGAAATTAATTTCAATTTCAAGG 8880
 QY 8881 GAAACAAATCATTCATTAAGATTAAGCAGAAATTTGTTGAACCGGAAGTAAATGTA 8940
 DB 8881 GAAACAAATCATTCATTAAGATTAAGCAGAAATTTGTTGAACCGGAAGTAAATGTA 8940
 QY 8941 GATGCTTTGAATGAACCTTGATGATGATGTAACAGAGAAAGTGGTGGTGGTGAAT 9000
 DB 8941 GATGCTTTGAATGAACCTTGATGATGATGTAACAGAGAAAGTGGTGGTGGTGAAT 9000
 QY 9001 GGTATTGGAAGATGATGATTAATGATGATTAAGAAAAATGTAAGCCAAATCCGA 9060

1021 GGAAGAATGTGACATTACAGCTGAAGCAAGAATTTCTATGATGCAACTTTAGTTACT 1080
1021 GGAAGAATGTGACATTACAGCTGAAGCAAGAATTTCTATGATGCAACTTTAGTTACT 1080
1081 AAGCTTGCAAGCACTCTTTAGCTTTGTTACAGGTTCTATTTCTCTATCAATTTTAAT 1140
1081 AAGCTTGCAAGCACTCTTTAGCTTTGTTACAGGTTCTATTTCTCTATCAATTTTAAT 1140
1141 GGAATTTTAAAGTTTATGACAAAGTAAGTCCAGTGTCTGTTATGGAAAAAGATGCCAAATGC 1200
1141 GGAATTTTAAAGTTTATGACAAAGTAAGTCCAGTGTCTGTTATGGAAAAAGATGCCAAATGC 1200
1201 GAAGCAACAGAAAGGAAGCAAAATATTCATTCTTACAGTGAAGTAAGCAACTATGSGA 1260
1201 GAAGCAACAGAAAGGAAGCAAAATATTCATTCTTACAGTGAAGTAAGCAACTATGSGA 1260
1261 GCAGCTACTTCTCCATTAAAAATTACCAATTTATTTGAGAAAAGCCAAATGAAAACCTT 1320
1261 GCAGCTACTTCTCCATTAAAAATTACCAATTTATTTGAGAAAAGCCAAATGAAAACCTT 1320
1321 CTCGATTCGAGCGGGGATATTTCTGCAAAAAGTAATTCGCAATGTAATTTGAAAGGA 1380
1321 CTCGATTCGAGCGGGGATATTTCTGCAAAAAGTAATTCGCAATGTAATTTGAAAGGA 1380
1381 GAAGTAAATGGAAGGGAAGCAAGATATTACTTCAAAATCGAAAATTAATTAATGATGCT 1440
1381 GAAGTAAATGGAAGGGAAGCAAGATATTACTTCAAAATCGAAAATTAATTAATGATGCT 1440
1441 TCTGTTCTGTTGGAACGATGAGAGATTCCAATAAGTAGCTCTTTCAGTATTTGGTGACG 1500
1441 TCTGTTCTGTTGGAACGATGAGAGATTCCAATAAGTAGCTCTTTCAGTATTTGGTGACG 1500
1501 GAAGAGAAAATTAATCTTCCGTCAAGATTCGTAAGAGCAAAAGTAAGTAAGCAAAAG 1560
1501 GAAGAGAAAATTAATCTTCCGTCAAGATTCGTAAGAGCAAAAGTAAGTAAGCAAAAG 1560
1561 GATGATGTAATGGAAGGGAAGCAAGTAATTCATTGAGCTGCTGTAAGAGTGA 1620
1561 GATGATGTAATGGAAGGGAAGCAAGTAATTCATTGAGCTGCTGTAAGAGTGA 1620
1621 TTGGGGGATAGTGTGATGAGGTTGGCTGCAAAATTTTCACTATATGCTTCTCTCC 1680
1621 TTGGGGGATAGTGTGATGAGGTTGGCTGCAAAATTTTCACTATATGCTTCTCTCC 1680
1681 CGTATGATGTAATGATGATATCTACATGCAAGAGCACTAAATGTAAGGCTCATAC 1740
1681 CGTATGATGTAATGATGATATCTACATGCAAGAGCACTAAATGTAAGGCTCATAC 1740
1741 ATTACTAAAAATAGGTTCTGCAAAACAGGATCTGATTTGGGAATTTCCAAAGTTTATGAA 1800
1741 ATTACTAAAAATAGGTTCTGCAAAACAGGATCTGATTTGGGAATTTCCAAAGTTTATGAA 1800
1801 GATCAGTTTATGATGATGATGATCTGCAAAACAGGATCTGATTTGGGAATTTCCAAAGTTTATGAA 1860
1801 GATCAGTTTATGATGATGATGATCTGCAAAACAGGATCTGATTTGGGAATTTCCAAAGTTTATGAA 1860
1861 GGAGAGACAGTGTCAATGAGGAAATTAAGTAAGCTTATTTAGTGTGCT 1920
1861 GGAGAGACAGTGTCAATGAGGAAATTAAGTAAGCTTATTTAGTGTGCT 1920
1921 GTGTCTGCAACCATGCAATATATATTTCTGCTTGTGCAATGGAAGAGTGA 1980
1921 GTGTCTGCAACCATGCAATATATATTTCTGCTTGTGCAATGGAAGAGTGA 1980
1981 AGACTTTCTTCAAGAGTGAAGGAGTAATGTAAGGCACTTAAATGAACTCAAAATCTT 2040
1981 AGACTTTCTTCAAGAGTGAAGGAGTAATGTAAGGCACTTAAATGAACTCAAAATCTT 2040
2041 CGAGGCACTTACGTAAGTGAAGTGTGCTGTACGAAGAGAAAGAAAAGAACTTAT 2100
2041 CGAGGCACTTACGTAAGTGAAGTGTGCTGTACGAAGAGAAAGAAAAGAACTTAT 2100
2101 GGAATGCAAGCTTTTATGAGAAATTAATTAATGCTTCTGTGCAATTCGCAAT 2160

2101 GGAATGCAAGCTTTTATGAAACTATTAATTAATGCTTCTGTGCAATTCGCAAT 2160
2161 CATGCTGAATTTGATGGAAGGAAATTTGATGTAATGCAACGTGAAATTAATTAAT 2220
2161 CATGCTGAATTTGATGGAAGGAAATTTGATGTAATGCAACGTGAAATTAATTAAT 2220
2221 AAAAATCTTCAAAATGCGCAAGTCTGTTATGTAATTAATTAATTAATTAATTAAT 2280
2221 AAAAATCTTCAAAATGCGCAAGTCTGTTATGTAATTAATTAATTAATTAATTAAT 2280
2281 TTTGAAAAGAAAGCAAACTCCAGAAATGATCGAAAGATTAATTAATTAATTAAT 2340
2281 TTTGAAAAGAAAGCAAACTCCAGAAATGATCGAAAGATTAATTAATTAATTAAT 2340
2341 TTATGTAATGATTTTCAAGAAATTTGATGGAAGAAACCGAGCTTTTATTAATGATGA 2400
2341 TTATGTAATGATTTTCAAGAAATTTGATGGAAGAAACCGAGCTTTTATTAATGATGA 2400
2401 AGAATGCAATTAATTTCTCCGATGGAACCTCAAAACAGGAACGCTATAGAAATGCA 2460
2401 AGAATGCAATTAATTTCTCCGATGGAACCTCAAAACAGGAACGCTATAGAAATGCA 2460
2461 AACTATGTTCAAGAGAAATGAAAAATTAAGAGAAAATTAACGAAAGATTTAAAGCT 2520
2461 AACTATGTTCAAGAGAAATGAAAAATTAAGAGAAAATTAACGAAAGATTTAAAGCT 2520
2521 TTTTCAAGAGATTAAGTGAACGTAATTAAGAACTTTGAATTTTACAGAGTGAAT 2580
2521 TTTTCAAGAGATTAAGTGAACGTAATTAAGAACTTTGAATTTTACAGAGTGAAT 2580
2581 TATGCAATTTTCAACCTTTTACCTCTCCGAGCTATGGAAGAAAGATTTCTCT 2640
2581 TATGCAATTTTCAACCTTTTACCTCTCCGAGCTATGGAAGAAAGATTTCTCT 2640
2641 GTGGAGAGAGCTGTTTCTGCTGATGAACAGAGATTAATTAAGTAATCCGTTGAAA 2700
2641 GTGGAGAGAGCTGTTTCTGCTGATGAACAGAGATTAATTAAGTAATCCGTTGAAA 2700
2701 GAGCTAACTTCTGCAAAAAAGTTTAAATTAATTAATTAATTAATTAATTAATTAAT 2760
2701 GAGCTAACTTCTGCAAAAAAGTTTAAATTAATTAATTAATTAATTAATTAATTAAT 2760
2761 GTGAATTAATGTTGAATTAATGGAATTTGCACTTGGAGTAATGCAAGATTAATTAAT 2820
2761 GTGAATTAATGTTGAATTAATGGAATTTGCACTTGGAGTAATGCAAGATTAATTAAT 2820
2821 GGAAGATTAATGTTCAAAAGATCGAAAAATTCAGCTATCTGTAAGAGCTTAAGAAAAGCT 2880
2821 GGAAGATTAATGTTCAAAAGATCGAAAAATTCAGCTATCTGTAAGAGCTTAAGAAAAGCT 2880
2881 GAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
2881 GAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
2941 TCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
2941 TCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
3001 ATCAGTAAGGCAAGGATTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
3001 ATCAGTAAGGCAAGGATTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
3061 GCTTTAAACAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
3061 GCTTTAAACAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
3121 ACGAAAAATGCGCGCTGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
3121 ACGAAAAATGCGCGCTGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
3181 GCTTCATTTGAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
3181 GCTTCATTTGAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240

Db 3181 GCTTCATTGAAGATTAATGACGAAAGACAAATTAATATGATTAAGAAATTAAGATGATGAA 3240
Oy 3241 GTAACAGTAACCTGCGAATCTTTAGAGTAGATGCAAAAACGACCGGAACATCAACAGT 3300
Db 3241 GTAACAGTAACCTGCGAATCTTTAGAGTAGATGCAAAAACGACCGGAACATCAACAGT 3300
Oy 3301 ATTTCTGTTGCCGAGGAATTAATAAGTTGGAAGTAAACCGAGTGAAGAAAACCGAAA 3360
Db 3301 ATTTCTGTTGCCGAGGAATTAATAAGTTGGAAGTAAACCGAGTGAAGAAAACCGAAA 3360
Oy 3361 TCAGAAAGAAAGACCGAGAGATTTTGTGCAAAATCGGAAAACAAAGTGAATCTGTATAA 3420
Db 3361 TCAGAAAGAAAGACCGAGAGATTTTGTGCAAAATCGGAAAACAAAGTGAATCTGTATAA 3420
Oy 3421 AATAAAATTACGAGATGATGATTCATTAAACAGAAAAAATTAACAATTACTTTCTGAA 3480
Db 3421 AATAAAATTACGAGATGATGATTCATTAAACAGAAAAAATTAACAATTACTTTCTGAA 3480
Oy 3481 GGAAGTAAAAAAGCGGGAACTCTTCGAAACGTTTCTCATCTCCGATPAAAGACCG 3540
Db 3481 GGAAGTAAAAAAGCGGGAACTCTTCGAAACGTTTCTCATCTCCGATPAAAGACCG 3540
Oy 3541 TCTTCAATTTGGAGCTTCTGGAAGTGTCTTTCAATAATATTAAGAAAACATCT 3600
Db 3541 TCTTCAATTTGGAGCTTCTGGAAGTGTCTTTCAATAATATTAAGAAAACATCT 3600
Oy 3601 GCTGTCATGATGAGTAAAGATTAATTGAAGGAGCAAAATPAAAGGTGAAGTACT 3660
Db 3601 GCTGTCATGATGAGTAAAGATTAATTGAAGGAGCAAAATPAAAGGTGAAGTACT 3660
Oy 3661 TCTTCTGATTTCACTTTTGTGAGACATGGGGCGGATCTGTCATCTCAATGGAATCAT 3720
Db 3661 TCTTCTGATTTCACTTTTGTGAGACATGGGGCGGATCTGTCATCTCAATGGAATCAT 3720
Oy 3721 ATTGGAATGGAATAAGCAATCACTGCTGTGTTTACTGAGAGCGGCTGCTGTAATAT 3780
Db 3721 ATTGGAATGGAATAAGCAATCACTGCTGTGTTTACTGAGAGCGGCTGCTGTAATAT 3780
Oy 3781 ATTCAAGATTAATAACAAGTCTTTGGTAAATAATAGTATTAATTCGAAATGCCAATPAAATTT 3840
Db 3781 ATTCAAGATTAATAACAAGTCTTTGGTAAATAATAGTATTAATTCGAAATGCCAATPAAATTT 3840
Oy 3841 AAAGTAAATGCTTTGAGTGAAGAACTCAAGTAGCAGAGAGAGAGCTTTGGAAGCAGTT 3900
Db 3841 AAAGTAAATGCTTTGAGTGAAGAACTCAAGTAGCAGAGAGAGAGCTTTGGAAGCAGTT 3900
Oy 3901 AAAGAAATGGAAGCAAGAAAAATTATCTATTGGGAATCTTCTGCTTCTATCACTTA 3960
Db 3901 AAAGAAATGGAAGCAAGAAAAATTATCTATTGGGAATCTTCTGCTTCTATCACTTA 3960
Oy 3961 AAAGAAATGGAAGCAAGAAAAATTATCTATTGGGAATCTTCTGCTTCTATCACTTA 3960
Db 3961 AAAGAAATGGAAGCAAGAAAAATTATCTATTGGGAATCTTCTGCTTCTATCACTTA 3960
Oy 3961 GTGAACAATGAAGTTCTGCAAAATCAGAAAAATTAATACAGTAGCAGAGAAATCTGAAAGC 4020
Db 3961 GTGAACAATGAAGTTCTGCAAAATCAGAAAAATTAATACAGTAGCAGAGAAATCTGAAAGC 4020
Oy 4021 CAAAAAATGATGATGATGTCATCTGCTTATCAAGCGAACACCAAGTAGCAGAGCTTTA 4080
Db 4021 CAAAAAATGATGATGATGTCATCTGCTTATCAAGCGAACACCAAGTAGCAGAGCTTTA 4080
Oy 4081 AATTTACAGAGCTGGAAGATCAAAATGGAACCTGTAAGGGGCTACGTGACCTGTGCGCAATTA 4140
Db 4081 AATTTACAGAGCTGGAAGATCAAAATGGAACCTGTAAGGGGCTACGTGACCTGTGCGCAATTA 4140
Oy 4141 AACCAACAAGTAAATGCTTCTATAGTGGTGGAGATATPACTTAAGTTAACTGAGCGGAC 4200
Db 4141 AACCAACAAGTAAATGCTTCTATAGTGGTGGAGATATPACTTAAGTTAACTGAGCGGAC 4200
Oy 4201 GCAAAAAGCTTTTGAACAACCACTCAAGTACTGTCGACGTGACGACGAGAGGACAATT 4260
Db 4201 GCAAAAAGCTTTTGAACAACCACTCAAGTACTGTCGACGTGACGAGAGGAGCAATT 4260
Oy 4261 AGTTCTGAGAGCGGATTAAGAAATTAATCAAGGGGCTGTTTCTGTCAATTAAGATTGACAT 4320
Db 4261 AGTTCTGAGAGCGGATTAAGAAATTAATCAAGGGGCTGTTTCTGTCAATTAAGATTGACAT 4320

Oy 4321 GACGTGAAGCTAGCGTTGATTAATCTTCCATCGAAGAGCTAAATGAATCAATGTCATT 4380
Db 4321 GACGTGAAGCTAGCGTTGATTAATCTTCCATCGAAGAGCTAAATGAATCAATGTCATT 4380
Oy 4381 GCCAAAGATGTCAAAGAGAGTTCTGATCTAGCAAAAAGAAATATACAGGCTTTACTAAATGA 4440
Db 4381 GCCAAAGATGTCAAAGAGAGTTCTGATCTAGCAAAAAGAAATATACAGGCTTTACTAAATGA 4440
Oy 4441 AAAGATTAATAATTTAAGAAATCGTGTATTAATACAGTGAAGAAATGTTATTAATACG 4500
Db 4441 AAAGATTAATAATTTAAGAAATCGTGTATTAATACAGTGAAGAAATGTTATTAATACG 4500
Oy 4501 AAGGAACAATAGAAAAAGCAAAAGAAAAAGAAAGAGACCGCTCATTTGTAATGCTGCTTTA 4560
Db 4501 AAGGAACAATAGAAAAAGCAAAAGAAAAAGAAAGAGACCGCTCATTTGTAATGCTGCTTTA 4560
Oy 4561 TCGGTTCTGGAACGATTAATCCGCTGAGAGAGTACCTATTTGCAATCACTGTTAA 4620
Db 4561 TCGGTTCTGGAACGATTAATCCGCTGAGAGAGTACCTATTTGCAATCACTGTTAA 4620
Oy 4621 AATAAATTTAAGAGAAATTAAGTGAAGCAATTAAGAAAGCGGAGAGGATPAAATTCAT 4680
Db 4621 AATAAATTTAAGAGAAATTAAGTGAAGCAATTAAGAAAGCGGAGAGGATPAAATTCAT 4680
Oy 4681 GCGAAACATGTAATGAGAGCAAAATCATCTACTGTGTGTAATGCGGCTTCTGGA 4740
Db 4681 GCGAAACATGTAATGAGAGCAAAATCATCTACTGTGTGTAATGCGGCTTCTGGA 4740
Oy 4741 CTGCTATACAGAAAGATGCTTTTTCAGAAATGGGATCTGAGACATGGAACATCTATCA 4800
Db 4741 CTGCTATACAGAAAGATGCTTTTTCAGAAATGGGATCTGAGACATGGAACATCTATCA 4800
Oy 4801 AATGACAGATGGAAGAGTGAATTAAGAAAGAAATTTCTGCTGATTCCTTAAATGTGAC 4860
Db 4801 AATGACAGATGGAAGAGTGAATTAAGAAAGAAATTTCTGCTGATTCCTTAAATGTGAC 4860
Oy 4861 GCAATAATTTCAATCTTTGGGGTGAATGTTGCGGGAACCAATGCGGTTCTCTTTCTACG 4920
Db 4861 GCAATAATTTCAATCTTTGGGGTGAATGTTGCGGGAACCAATGCGGTTCTCTTTCTACG 4920
Oy 4921 GCGGTAGAGCTGCTTTTGCAGATTAATACCTCTTCAATAATAAACCTCTGCTTGAATTA 4980
Db 4921 GCGGTAGAGCTGCTTTTGCAGATTAATACCTCTTCAATAATAAACCTCTGCTTGAATTA 4980
Oy 4981 GGAACGAAGGTAAATCCTTTTAAAGTGAAGAAATGCAAAAGTCAATGTACAAAGCTTTGAAAT 5040
Db 4981 GGAACGAAGGTAAATCCTTTTAAAGTGAAGAAATGCAAAAGTCAATGTACAAAGCTTTGAAAT 5040
Oy 5041 GATTCATATTTACAAACGTTTCTGCTGAGAGCGCTCAAGTATTAAGCAGAGCTGGAATC 5100
Db 5041 GATTCATATTTACAAACGTTTCTGCTGAGAGCGCTCAAGTATTAAGCAGAGCTGGAATC 5100
Oy 5101 GGAAGAAATGATCTGTCAATCGTGTCTGATGAAAACGGAAGCTTTAGTTAGTATCT 5160
Db 5101 GGAAGAAATGATCTGTCAATCGTGTCTGATGAAAACGGAAGCTTTAGTTAGTATCT 5160
Oy 5161 GAGTTTGAAGAGTAAGTCTTTCAATGTATGATCAAAAGTCAAAAGTCAATTAATCA 5220
Db 5161 GAGTTTGAAGAGTAAGTCTTTCAATGTATGATCAAAAGTCAAAAGTCAATTAATCA 5220
Oy 5221 ATTGCCGGAATGCAAAATGAGAAAAAGCGGCTGAGATGAGCAACGTTGCTCATACA 5280
Db 5221 ATTGCCGGAATGCAAAATGAGAAAAAGCGGCTGAGATGAGCAACGTTGCTCATACA 5280
Oy 5281 AATAATTGAAAAACAATCAGTTATAGCTATTGTAAAAAACAAGTAAATTAACAACGCGAAT 5340
Db 5281 AATAATTGAAAAACAATCAGTTATAGCTATTGTAAAAAACAAGTAAATTAACAACGCGAAT 5340
Oy 5341 GATCAAGATAGAAAAATATCAATGTGACCTGCAAAAAGATTAATCAATGACCAATATATA 5400
Db 5341 GATCAAGATAGAAAAATATCAATGTGACCTGCAAAAAGATTAATCAATGACCAATATATA 5400

QY	5401	GCAGTCGGAAGTTGGAGGACAAAGAGCCTCTGTGCAGAGAGCTTCTGCAGTACTACC	5460
Db	5401	GCAGTCGGAAGTTGGAGGACAAAGAGCCTCTGTGCAGAGAGCTTCTGCAGTACTACC	5460
QY	5461	TTGGAATAAGACAGTTTCTTCATGTTGATCAACATCGATATTTGACAAAGATTTTGAAGAA	5520
Db	5461	TTGGAATAAGACAGTTTCTTCATGTTGATCAACATCGATATTTGACAAAGATTTTGAAGAA	5520
QY	5521	GAATAATATGAAATAAAGAAAAGCCAAATGTTTAACTTACCTGMAAAATGACAGTCAA	5580
Db	5521	GAATAATATGAAATAAAGAAAAGCCAAATGTTTAACTTACCTGMAAAATGACAGTCAA	5580
QY	5581	GTGTCACAAATGCGACAGTGTCTTCCGAGACAAATGTGACAAAGTGCAGATGAGAGCTGGA	5640
Db	5581	GTGTCACAAATGCGACAGTGTCTTCCGAGACAAAGTGCAGATGAGAGCTGGA	5640
QY	5641	GTAGCACTTAAATAAATTAACAATACTTCCGACATATTAATAAATAATGACTCAAAAT	5700
Db	5641	GTAGCACTTAAATAAATTAACAATACTTCCGACATATTAATAAATAATGACTCAAAAT	5700
QY	5701	GTAGCAATGCTTGGTGTAAAAAGCAAAATCTCATCTATTTAAAACAATTGGAATTGGA	5760
Db	5701	GTAGCAATGCTTGGTGTAAAAAGCAAAATCTCATCTATTTAAAACAATTGGAATTGGA	5760
QY	5761	GCTGAGTTGGAGCTGAGAGAGCTGAGATGACAGTTCTGTAGCAGTGAATAAGATTGTA	5820
Db	5761	GCTGAGTTGGAGCTGAGAGAGCTGAGATGACAGTTCTGTAGCAGTGAATAAGATTGTA	5820
QY	5821	AATAATACGATAGACGAATTTAAATCAGTGCAGAAAAATCATCTCGAAGGGAAATGTCCGAGTT	5880
Db	5821	AATAATACGATAGACGAATTTAAATCAGTGCAGAAAAATCATCTCGAAGGGAAATGTCCGAGTT	5880
QY	5881	ATTACAGAGTCTGATGCGGTAAATTGCTAAATTATGACAGAAAACAGTGTCTGAGTGGCCCGT	5940
Db	5881	ATTACAGAGTCTGATGCGGTAAATTGCTAAATTATGACAGAAAACAGTGTCTGAGTGGCCCGT	5940
QY	5941	GCAGCAATAGGAGCCTCAACCAAGTGTGATGAATAATACAGAGTTCACAAAAGCATATGTA	6000
Db	5941	GCAGCAATAGGAGCCTCAACCAAGTGTGATGAATAATACAGAGTTCACAAAAGCATATGTA	6000
QY	6001	AAAATTTCTACAGTGAATTGCTTAAAGAAAGAAACAGATGATTATTACTACATCAAGGGCAA	6060
Db	6001	AAAATTTCTACAGTGAATTGCTTAAAGAAAGAAACAGATGATTATTACTACATCAAGGGCAA	6060
QY	6061	GTAGATAAAGGTGATGAATAAAGTTTCAAAATCTTAATATTTAAGAAAGTCTTATCACAA	6120
Db	6061	GTAGATAAAGGTGATGAATAAAGTTTCAAAATCTTAATATTTAAGAAAGTCTTATCACAA	6120
QY	6121	AAAAGAAAAATAGTAATTAATAAAGAAAGATTTGTACCAATAGTCAAGCTCATATCTTTA	6180
Db	6121	AAAAGAAAAATAGTAATTAATAAAGAAAGATTTGTACCAATAGTCAAGCTCATATCTTTA	6180
QY	6181	AAATTTTATTGGCAAATGCGCGTGTCTCAGACAAAGCCGAGTGGCAGAGAACTGTTAAT	6240
Db	6181	AAATTTTATTGGCAAATGCGCGTGTCTCAGACAAAGCCGAGTGGCAGAGAACTGTTAAT	6240
QY	6241	ATCAACAAGTTTATGAGAAACAGAACTCTTGTGAAAAATTTCTAATTTTAAATGCAAAA	6300
Db	6241	ATCAACAAGTTTATGAGAAACAGAACTCTTGTGAAAAATTTCTAATTTTAAATGCAAAA	6300
QY	6301	CATTATTCGTAAAAATCAGAGATTACAGAAATTCATCGAGATGATGAGTTCTGTGGT	6360
Db	6301	CATTATTCGTAAAAATCAGAGATTACAGAAATTCATCGAGATGATGAGTTCTGTGGT	6360
QY	6361	GTTTGTGGAATGTAGAGATGAGAGCTTCTCTGATACCAATATTTAATAAAGAAATACC	6420
Db	6361	GTTTGTGGAATGTAGAGATGAGAGCTTCTCTGATACCAATATTTAATAAAGAAATACC	6420
QY	6421	AAGACAAGATGTGAAAAATCTACATGTCTGATGAAGGTTTCCGAGAAAGAGCTGAATTT	6480
Db	6421	AAGACAAGATGTGAAAAATCTACATGTCTGATGAAGGTTTCCGAGAAAGAGCTGAATTT	6480
QY	6481	ACAGCAGATTTCTTAACAGGAATTTCTCTTTTGGAGTCCGAGATTCGACAGCCGGGGTA	6540

Db	6481	ACACGACGATTTCTAAGCAGGAA	TTTCTCTTTTGAGATCGGAGTCCGACGACCGGGGTA	6540
Qy	6541	GGAGCCGAGTGCAGGAA	CGTTTCCGTAATCAATTTGCAGGAAGACGGAATGAT	6600
Db	6541	GGAGCCGAGTGGCAGGAA	CCGTTTCGTAAATCAATTTGCAGGAAGACGGAATGAT	6600
Qy	6601	GTGGAGAGACCAAGATTTTGGTAAAA	AAAGCTGAGTTTCAGCAAAACGTTATAGTTCT	6660
Db	6601	GTGGAGAGACCAAGATTTTGGTAAAA	AAAGCTGAGTTTCAGCAAAACGTTATAGTTCT	6660
Qy	6661	GTTCGCAATTGGAATGCGCAGTCCGAGTGGAGTGGCTCAAA	GGAGCTGTGAATTGGACGACA	6720
Db	6661	GTTCGCAATTGGAATGCGCAGTCCGAGTGGAGTGGCTCAAA	GGAGCTGTGAATTGGACGACA	6720
Qy	6721	GTGGCAGTTACCAAGAATGATCAAA	CACGAGCAAGAGTGAATAATTCATAAATTATG	6780
Db	6721	GTGGCAGTTACCAAGAATGATCAAA	CACGAGCAAGAGTGAATAATTCATAAATTATG	6780
Qy	6781	ACTGGAACCAAGTTAGATGTATATGCA	GAATAAGATTAATAATAGGTACTGGAAATCGCT	6840
Db	6781	ACTGGAACCAAGTTAGATGTATATGCA	GAATAAGATTAATAATAGGTACTGGAAATCGCT	6840
Qy	6841	TCAGCCGAGCTGGAAATTCCTTCGACCGCGAGTACTGGA	GGTTCCTGCAATATATTT	6900
Db	6841	TCAGCCGAGCTGGAAATTCCTTCGACCGCGAGTACTGGA	GGTTCCTGCAATATATTT	6900
Qy	6901	GCAATATAGGTAGAAACAGATATGATCATATGATCTTTAC	CTCTTCACTGAGTAAAT	6960
Db	6901	GCAATATAGGTAGAAACAGATATGATCATATGATCTTTAC	CTCTTCACTGAGTAAAT	6960
Qy	6961	GTAAAGCTCTTAATAAATTTGCAATTCCTTGACAGCCG	TGAGAGACCGCAGCTCTT	7020
Db	6961	GTAAAGCTCTTAATAAATTTGCAATTCCTTGACAGCCG	TGAGAGACCGCAGCTCTT	7020
Qy	7021	GCACAGTTACCGGAGTGTCTTGTTAA	CACTAATATGTTCTGTATATGCTCGACTT	7080
Db	7021	GCACAGTTACCGGAGTGTCTTGTTAA	CACTAATATGTTCTGTATATGCTCGACTT	7080
Qy	7081	CACAATACTGTGATTTGACTTCCGTACGAGAA	AAAGTATGTAAACGGCAAAAGAGAA	7140
Db	7081	CACAATACTGTGATTTGACTTCCGTACGAGAA	AAAGTATGTAAACGGCAAAAGAGAA	7140
Qy	7141	AAAAATATTAAGCAAAACGACGAAATGCA	GAAATCCGAGGACGCAATCGAGCCAT	7200
Db	7141	AAAAATATTAAGCAAAACGACGAAATGCA	GAAATCCGAGGACGCAATCGAGCCAT	7200
Qy	7201	GCTTGTGTAATTAATTTTGGAA	CAGCTGTAGAGATGCAAAAAATTCGAAAGAAAGAA	7260
Db	7201	GCTTGTGTAATTAATTTTGGAA	CAGCTGTAGAGATGCAAAAAATTCGAAAGAAAGAA	7260
Qy	7261	ACAGAGTTTAAATAAATTTTGGAA	CGAATTAACAAAGAACAGATAAAAAGTAAATGAT	7320
Db	7261	ACAGAGTTTAAATAAATTTTGGAA	CGAATTAACAAAGAACAGATAAAAAGTAAATGAT	7320
Qy	7321	GCTACGAAAAAAATCTTCAATTCAGAGGTATTTCTA	CAGAAAGATTACTTCTGTAAACGG	7380
Db	7321	GCTACGAAAAAAATCTTCAATTCAGAGGTATTTCTA	CAGAAAGATTACTTCTGTAAACGG	7380
Qy	7381	GATAGAGAGATTACTCAGGAGAAAGAA	TTAAAGCCATTGTGAGACTTCTGATATTAAT	7440
Db	7381	GATAGAGAGATTACTCAGGAGAAAGAA	TTAAAGCCATTGTGAGACTTCTGATATTAAT	7440
Qy	7441	GGAAAAAATGTGAGATTAACAACAGAGA	CAAGAAATATATCACTTCACTGAGTGGTATTTG	7500
Db	7441	GGAAAAAATGTGAGATTAACAACAGAGA	CAAGAAATATATCACTTCACTGAGTGGTATTTG	7500
Qy	7501	GGAACTGACAGTCTTGTCCGCAATCA	GAAACAGTGGCAGTTACAAATATTAAGAAAT	7560
Db	7501	GGAACTGACAGTCTTGTCCGCAATCA	GAAACAGTGGCAGTTACAAATATTAAGAAAT	7560
Qy	7561	TCGGAGTTACTGTGAAATTTCTTTGTGA	AGACGTGAATAATGTTATGATCG	7620

Db 7561 TCCGAGTTACTGTTGAAAATTCITTTGTGAAAGCAGCTGAAAAAGTAATGTAGATCG 7620
Qy 7621 GATATTACAGGAAATGTTGCTTTAAACAGCATATCAAGGTCCTGTAGAGCATTTGGGAATA 7680
Db 7621 GATATTACAGGAAATGTTGCTTTAAACAGCATATCAAGGTCCTGTAGAGCATTTGGGAATA 7680
Qy 7661 GAGAGTCCCTATGCAAAATTAAATCTTAATGAGAGATCAAAATATCAGTATTTAAAAATCT 7740
Db 7661 GAGAGTCCCTATGCAAAATTAAATCTTAATGAGAGATCAAAATATCAGTATTTAAAAATCT 7740
Qy 7741 AAGCTATTAGGAAAAAATATGATGTTAATTGTAATGAATGTAATCCGAAATTGAGCCGAA 7800
Db 7741 AAGCTATTAGGAAAAAATATGATGTTAATTGTAATGAATGTAATCCGAAATTGAGCCGAA 7800
Qy 7801 GCAAAAAGATTAAACCGTAGAGAGCGGTAGCTCCGAGCCATTATCTCAAAAAGCAAGAAT 7860
Db 7801 GCAAAAAGATTAAACCGTAGAGAGCGGTAGCTCCGAGCCATTATCTCAAAAAGCAAGAAT 7860
Qy 7861 GAAATGAAATTCAGAGGTTGAAAATTGAGAAAGTATTTTCAATGAAGAAATAGAGTAACT 7920
Db 7861 GAAATGAAATTCAGAGGTTGAAAATTGAGAAAGTATTTTCAATGAAGAAATAGAGTAACT 7920
Qy 7921 AGCCCTTTTAAAGAAATTGGAAGAGAAATCAATGTCAAAGTGGAAAAAGCAAGATG 7980
Db 7921 AGCCCTTTTAAAGAAATTGGAAGAGAAATCAATGTCAAAGTGGAAAAAGCAAGATG 7980
Qy 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGAGGCGCAGAAATTTTCCGAA 8040
Db 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGAGGCGCAGAAATTTTCCGAA 8040
Qy 8041 GCAAAAAGATGCCGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100
Db 8041 GCAAAAAGATGCCGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100
Qy 8101 CATGAGATTAATGTGAATATGGAAGCAACATTAATAAGAAATGAAACAGCACTTTCTAA 8160
Db 8101 CATGAGATTAATGTGAATATGGAAGCAACATTAATAAGAAATGAAACAGCACTTTCTAA 8160
Qy 8161 GCAGTAACAGGTTCTGTATTGGGAGAGTGGAGTCAACAAGGCGAAGCTACTGCTCA 8220
Db 8161 GCAGTAACAGGTTCTGTATTGGGAGAGTGGAGTCAACAAGGCGAAGCTACTGCTCA 8220
Qy 8221 GGTAAACTATGTGTGAAGTTGAGAGAGAAATTTGTCAGAACAATCGATGTAATCA 8280
Db 8221 GGTAAACTATGTGTGAAGTTGAGAGAGAAATTTGTCAGAACAATCGATGTAATCA 8280
Qy 8281 ATTTCTAAGTAGAAGTTTGGATGAATGAATGAATCTGCTAAATCTTCTGTAGTATCA 8340
Db 8281 ATTTCTAAGTAGAAGTTTGGATGAATGAATGAATCTGCTAAATCTTCTGTAGTATCA 8340
Qy 8341 GGAATGAGAGAGAAATGCGGAGCAGAGATGAATCTTCTACAGCACAAAAGTAACT 8400
Db 8341 GGAATGAGAGAGAAATGCGGAGCAGAGATGAATCTTCTACAGCACAAAAGTAACT 8400
Qy 8401 GAATCCGTAAGTTCTTTACGAAAGCAGATTAATGAATGAATGATTAACAAAAAATAT 8460
Db 8401 GAATCCGTAAGTTCTTTACGAAAGCAGATTAATGAATGAATGATTAACAAAAAATAT 8460
Qy 8461 ATTTCAAGAGTCAATGCTCTCTCTTTAAATGATCAAAAGATGAAGCAATATAGATCT 8520
Db 8461 ATTTCAAGAGTCAATGCTCTCTCTTTAAATGATCAAAAGATGAAGCAATATAGATCT 8520
Qy 8521 TTAGCGGTAGCCGGTGTGCATGCACAAAGAACAAAGCAATTAGAGATCAAAAG 8580
Db 8521 TTAGCGGTAGCCGGTGTGCATGCACAAAGAACAAAGCAATTAGAGATCAAAAG 8580
Qy 8581 TTAACCTTACAACTGTAAATGAGAGAAACGTATCTCAACTTCGTGCAAAAAGCTTTGCT 8640
Db 8581 TTAACCTTACAACTGTAAATGAGAGAAACGTATCTCAACTTCGTGCAAAAAGCTTTGCT 8640
Qy 8641 AAAAAATGAATTAATGAGAAATGTAAGAAATGGAAGAGAGAGCTTGTGAGCGGAAACA 8700
Db 8641 AAAAAATGAATTAATGAGAAATGTAAGAAATGGAAGAGAGCTTGTGAGCGGAAACA 8700

Qy 8701 GCAGCCGTTGAAAAATTATACAAAGATACTACAGAGCATTTGTCAGAGAAATTTGGGAA 8760
Db 8701 GCAGCCGTTGAAAAATTATACAAAGATACTACAGAGCATTTGTCAGAGAAATTTGGGAA 8760
Qy 8761 ATTGAGATTAATTAGAAAGCAATTTGCAAGAGATTAATCAGATTTGAGAGTCAACGAGAC 8820
Db 8761 ATTGAGATTAATTAGAAAGCAATTTGCAAGAGATTAATCAGATTTGAGAGTCAACGAGAC 8820
Qy 8821 GGAACCAAGAGAGGTCCTGTGCGAAAGATGCTTTCTGTGAAAAATACAAATTTCAAGG 8880
Db 8821 GGAACCAAGAGAGGTCCTGTGCGAAAGATGCTTTCTGTGAAAAATACAAATTTCAAGG 8880
Qy 8881 GAAACAAAATCATCATTTGAAAGATTAAGCCAGAAATTTGGAACCGGAAGTGAATGTA 8940
Db 8881 GAAACAAAATCATCATTTGAAAGATTAAGCCAGAAATTTGGAACCGGAAGTGAATGTA 8940
Qy 8941 GATGCTTTGAATGAATCTGATGATGATCTACAGAGAAAGAGTGTGCTATGTGAAAT 9000
Db 8941 GATGCTTTGAATGAATCTGATGATGATCTACAGAGAAAGAGTGTGCTATGTGAAAT 9000
Qy 9001 GGTATTGGAATGTTGATGTTAATGTAATTAAGAAAAATGTAAGCCAAAATCGGA 9060
Db 9001 GGTATTGGAATGTTGATGTTAATGTAATTAAGAAAAATGTAAGCCAAAATCGGA 9060
Qy 9061 AGACATGCTATTGTAGAAACTACAGAAACAAAGATATCAAGCATTTACAAAGCAAAA 9120
Db 9061 AGACATGCTATTGTAGAAACTACAGAAACAAAGATATCAAGCATTTACAAAGCAAAA 9120
Qy 9121 GTAATATTTCTGGAAGAAAGAGACGCTGACGCTCAGCTCAGTCAATGCAATGTAACAAT 9180
Db 9121 GTAATATTTCTGGAAGAAAGAGAGACGCTGACGCTCAGCTCAGTCAATGTAACAAT 9180
Qy 9181 TCCAATGAGATGATTAATTAATAATTGGCAAGAGATGATCTTCTCAATTAATAACC 9240
Db 9181 TCCAATGAGATGATTAATTAATAATTGGCAAGAGATGATCTTCTCAATTAATAACC 9240
Qy 9241 AAAAAATCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9300
Db 9241 AAAAAATCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9300
Qy 9301 GGGGTGCTGAAGCAAGAGGTGCGAGAGCCAAACGACAGTATGTAAGAAATCAATA 9360
Db 9301 GGGGTGCTGAAGCAAGAGGTGCGAGAGCCAAACGACAGTATGTAAGAAATCAATA 9360
Qy 9361 AATGAACATAATATGTTGATTTAGCAGGAAAAATTAACAGAGGGAACATCAATGTA 9420
Db 9361 AATGAACATAATATGTTGATTTAGCAGGAAAAATTAACAGAGGGAACATCAATGTA 9420
Qy 9421 TATGCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9480
Db 9421 TATGCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9480
Qy 9481 GCAAAAAGTCAATGCTGCAAGCTGCTCGCAACTCCACTATTGAAAAAAGAAAGTAAAA 9540
Db 9481 GCAAAAAGTCAATGCTGCAAGCTGCTCGCAACTCCACTATTGAAAAAAGAAAGTAAAA 9540
Qy 9541 TTTAATTAATGGAATCCGAGAAATTTAAATAATCTGCGAAGATTTGGAAGGAAAGCTAAT 9600
Db 9541 TTTAATTAATGGAATCCGAGAAATTTAAATAATCTGCGAAGATTTGGAAGGAAAGCTAAT 9600
Qy 9601 AAAAAAGCTGCGTAGATCTTAATCAGGTAGACTGATACAGATTAATAATATACATGCGAT 9660
Db 9601 AAAAAAGCTGCGTAGATCTTAATCAGGTAGACTGATACAGATTAATAATATACATGCGAT 9660
Qy 9661 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAGAGAGAGAAAAAGGAAA 9720
Db 9661 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAGAGAGAGAAAAAGGAAA 9720
Qy 9721 AAATGA 9726
Db 9721 AAATGA 9726

Db 2774 ATTACTAAATAATAGGTTCTGCAAAACAGATCTGATTTGGAACTTCCAAAGTTTATGAT 2833
Qy 1801 GATCGATTTTGAATGCAAGTCACTAAATCAATTTAGATGCAATAAAACAGCGGTTT 1860
Db 2834 GATCGATTTTGAATGCAAGTCACTAAATCAATTTAGATGCAATAAAACAGCGGTTT 2893
Qy 1861 GAGAGAGACAGTGTCAATGAGGAAATTAAGATAAGCTAACGACTTATTTAGTCCGT 1920
Db 2894 GAGAGAGACAGTGTCAATGAGGAAATTAAGATAAGCTAACGACTTATTTAGTCCGT 2953
Qy 1921 GTGTCTGCACCATATGCAAACTATATAATTTCTGCTCTGTGGCAATGAGAGAGTGA 1980
Db 2954 GTGTCTGCACCATATGCAAACTATATAATTTCTGCTCTGTGGCAATGAGAGAGTGA 3013
Qy 1981 AGACTTTCTTCAGAGTGGAGGAGTAAATGTAAGGGCACTTAAATGAAGCTCAAAATCTT 2040
Db 3014 AGACTTTCTTCAGAGTGGAGGAGTAAATGTAAGGGCACTTAAATGAAGCTCAAAATCTT 3073
Qy 2041 CGAGCGACTACGTCAAGTGGAAAGTGTGCTGTACGAAAGGAGAAAGAAAGAACTTAT 2100
Db 3074 CGAGCGACTACGTCAAGTGGAAAGTGTGCTGTACGAAAGGAGAAAGAAAGAAAGCTTAT 3133
Qy 2101 GGAATTCGACGCGATTTTATGAAAATTAATGCTTCTGTGCAATGTCGCAAT 2160
Db 3134 GGAATTCGACGCGATTTTATGAAAATTAATGCTTCTGTGCAATGTCGCAAT 3193
Qy 2161 CATGCTGAATGGTATCCGAAAGAAATTTGATCAACAGTGAATTAATTTGAATAT 2220
Db 3194 CATGCTGAATGGTATCCGAAAGAAATTTGATCAACAGTGAATTAATTTGAATAT 3253
Qy 2221 AAAAAATCCTTCAAAAATGCAAAAGTCTGTATATGATAATTAAGAACTTTTAAAGAGCT 2280
Db 3254 AAAAAATCCTTCAAAAATGCAAAAGTCTGTATATGATAATTAAGAACTTTTAAAGAGCT 3313
Qy 2281 TTTGGAAAAAGAAACGAAATCTCAGAAATGATCCGAAAGATTTGAATCTATGAAAA 2340
Db 3314 TTTGGAAAAAGAAACGAAATCTCAGAAATGATCCGAAAGATTTGAATCTATGAAAA 3373
Qy 2341 TTATTTGAATGCAATTTTCAAAAAATTTGATGAGAAACCGAGCTTTTCTAAATGGTGA 2400
Db 3374 TTATTTGAATGCAATTTTCAAAAAATTTGATGAGAAACCGAGCTTTTCTAAATGGTGA 3433
Qy 2401 AGAATGACAATTAATCTTCCGATGAGAACTTCAAAAAACGAACTCTATAGAAATTTGCA 2460
Db 3434 AGAATGACAATTAATCTTCCGATGAGAACTTCAAAAAACGAACTCTATAGAAATTTGCA 3493
Qy 2461 AACTATGTTCAAGGAGAAATGAAAAATTTAGAGAAAAATTAACGAAAGATTTAAAGCT 2520
Db 3494 AACTATGTTCAAGGAGAAATGAAAAATTTAGAGAAAAATTAACGAAAGATTTAAAGCT 3553
Qy 2521 TTTTCAAGAGATGATGAGGACTGATTAAGAAACCTTGAATTTTACAGAGTATAGAAAT 2580
Db 3554 TTTTCAAGAGATGATGAGGACTGATTAAGAAACCTTGAATTTTACAGAGTATAGAAAT 3613
Qy 2581 TATGCAAAATTTTCAACTTTTACCTCTTCCGAGCTTAATGAGAAAGAGATGTTCTCT 2640
Db 3614 TATGCAAAATTTTCAACTTTTACCTCTTCCGAGCTTAATGAGAAAGAGATGTTCTCT 3673
Qy 2641 GTGGAGAGAGCTGTTTCTGTGGTGAACAAGGAAATTAATGCAAGTATCCGTTGAAAA 2700
Db 3674 GTGGAGAGAGCTGTTTCTGTGGTGAACAAGGAAATTAATGCAAGTATCCGTTGAAAA 3733
Qy 2701 GGAGCTAAACCTGCTGCAAAAAAAGATTAATATTAATAAGCTAATTAAGCAAAAAACA 2760
Db 3734 GGAGCTAAACCTGCTGCAAAAAAAGATTAATATTAATAAGCTAATTAAGCAAAAAACA 3793
Qy 2761 GTGAATTTAGTTGAAAAATTTGCACTTGCGAAGACAGTACATCCGAAAGTCCAGTCGA 2820
Db 3794 GTGAATTTAGTTGAAAAATTTGCACTTGCGAAGACAGTACATCCGAAAGTCCAGTCGA 3853
Qy 2821 GGAAGATTTAATGTTCAAAAGATCGAAAAATTCAGCTATCTGATGAAGCTTAAGAAAAAGCT 2880

Db 3854 GGAAGATTTAATGTTCAAAAGATCGAAAAATTCAGCTATCTGATGAAGCTTAAGAAAAAGCT 3913
Qy 2881 GAATTTATCAGAGAAAAATTAATTAATGCAATGCAATTTTCAATGTAAGCGGGA 2940
Db 3914 GAATTTATCAGAGAAAAATTAATTAATGCAATGCAATTTTCAATGTAAGCGGGA 3973
Qy 2941 TCTTTTAAATGCTGCTCAGGTGAGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGT 3000
Db 3974 TCTTTTAAATGCTGCTCAGGTGAGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGT 4033
Qy 3001 ATCAGTAAAGGCAAGATTTTCAATTTGATGACGACATATTTGAAAGCTAATTAATAAATT 3060
Db 4034 ATCAGTAAAGGCAAGATTTTCAATTTGATGACGACATATTTGAAAGCTAATTAATAAATT 4093
Qy 3061 GCTTTTAAACAGTAAAGATTAATGATCTTCTGTTGGAATGTCGCGGTTACGCGGAATCGGA 3120
Db 4094 GCTTTTAAACAGTAAAGATTAATGATCTTCTGTTGGAATGTCGCGGTTACGCGGAATCGGA 4153
Qy 3121 ACGAAAAATGCGCGGCTCGCGGTTGCTGTTGCGGTAATGATTAATTTCAAAACAA 3180
Db 4154 ACGAAAAATGCGCGGCTCGCGGTTGCTGTTGCGGTAATGATTAATTTCAAAACAA 4213
Qy 3181 GCTTCCATTTGAAGATTAATGACGAGAGCAAAATTAATGATTAAGATTAAGATGATA 3240
Db 4214 GCTTCCATTTGAAGATTAATGACGAGAGCAAAATTAATGATTAAGATTAAGATGATA 4273
Qy 3241 GTAACAGTAACTGCGGAATCTTTAGAAAGTATGCAAAAAACGACCGGAACATCAACAGT 3300
Db 4274 GTAACAGTAACTGCGGAATCTTTAGAAAGTATGCAAAAAACGACCGGAACATCAACAGT 4333
Qy 3301 ATTTCTGTTGCGGAGAAATTAATTAAGTTGAAAGTAAACGAGTGAAGAAAAACGAA 3360
Db 4334 ATTTCTGTTGCGGAGAAATTAATTAAGTTGAAAGTAAACGAGTGAAGAAAAACGAA 4393
Qy 3361 TCAGAAAAAGAACAGAGGAGTTTTTTGGCAAAATTCGAAACAAAGTGAATCTCTGTAAA 3420
Db 4394 TCAGAAAAAGAACAGAGGAGTTTTTTGGCAAAATTCGAAACAAAGTGAATCTCTGTAAA 4453
Qy 3421 AATTAATTAATGAGATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3480
Db 4454 AATTAATTAATGAGATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 4513
Qy 3481 GGAATTAATAAAAGCGGGAATCTTCTTCGAAAGTTTCTCATACCTCCGATTAAGAGACG 3540
Db 4514 GGAATTAATAAAAGCGGGAATCTTCTTCGAAAGTTTCTCATACCTCCGATTAAGAGACG 4573
Qy 3541 TCTTTCAAGTTGGAGGCTTCTGGAAGTGTCTTTCAATAATTAATAAAAGAAACATCT 3600
Db 4574 TCTTTCAAGTTGGAGGCTTCTGGAAGTGTCTTTCAATAATTAATAAAAGAAACATCT 4633
Qy 3601 GCTGTCGATGATGAGTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3660
Db 4634 GCTGTCGATGATGAGTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4693
Qy 3661 TCTTCTGATTTCTACTTTTGTGAGCAATGGGCGGAATCTGCTGCACTTCACTGGAATCAT 3720
Db 4694 TCTTCTGATTTCTACTTTTGTGAGCAATGGGCGGAATCTGCTGCACTTCACTGGAATCAT 4753
Qy 3721 ATTGGAAGTGAATTAAGCAATCATCAATGCTGTTTATGCTGAGAGGCTGCTGTAAATAT 3780
Db 4754 ATTGGAAGTGAATTAAGCAATCATCAATGCTGTTTATGCTGAGAGGCTGCTGTAAATAT 4813
Qy 3781 ATTCAAGATTAACAAGGCTTGTGTTTAAATAATGTAATTTGAAATGCAATTAATTTT 3840
Db 4814 ATTCAAGATTAACAAGGCTTGTGTTTAAATAATGTAATTTGAAATGCAATTAATTTT 4873
Qy 3841 AAAGTAAATGCTTTGAGTGAAGAACTCACTATGACAGAGAGCAAGTTTGAAGACAGTT 3900
Db 4874 AAAGTAAATGCTTTGAGTGAAGAACTCACTATGACAGAGAGCAAGTTTGAAGACAGTT 4933
Qy 3901 AAAGAAAGTGAAGCAAGAAAAAGTTATCTATTTGGAATCTTCTGCTTCTATCACTTA 3960
Db 4934 AAAGAAAGTGAAGCAAGAAAAAGTTATCTATTTGGAATCTTCTGCTTCTATCACTTA 4993

Qy	3661	GTGAACAATGAAGTTTCTGCAAAATCGAAAAATTAATCACTAGACAGAGAAATCTGAAGC	4020
Db	4994	GTGAACAATGAAGTTTCTGCAAAATCGAAAAATTAATCACTAGACAGAGAAATCTGAAGC	5053
Qy	4021	CAAAAAATGAGTGTGATGTCACTGCTTATCGACGGACACCCAACTGACAGAGCTTTA	4080
Db	5054	CAAAAAATGAGTGTGATGTCACTGCTTATCGACGGACACCCAACTGACAGAGCTTTA	5113
Qy	4081	AATTTACAAGCTGAAAAAGTCAAAATGGAATCTGTAGGGCTACTGTGACTGTTGCCAAATTA	4140
Db	5114	AATTTACAAGCTGAAAAAGTCAAAATGGAATCTGTAGGGCTACTGTGACTGTTGCCAAATTA	5173
Qy	4141	AACAACAAGTAATTCGTTCTATTAGTGTGGAGATTAATACTTAATCGAGCGAC	4200
Db	5174	AACAACAAGTAATTCGTTCTATTAGTGTGGAGATTAATACTTAATCGAGCGAC	5233
Qy	4201	GCAAAAGCTCTTTAGCAACACCTCAAGTGACTGCTGACGAGACAGAGGGAGCAATT	4260
Db	5234	GCAAAAGCTCTTTAGCAACACCTCAAGTGACTGCTGACGAGACAGAGGGAGCAATT	5293
Qy	4261	AGTCTGGAGGGGATTAAGAAATTATCAAGGGCTGTTCTGTCAATTAAGTTGCAAT	4320
Db	5294	AGTCTGGAGGGGATTAAGAAATTATCAAGGGCTGTTCTGTCAATTAAGTTGCAAT	5353
Qy	4321	GACGTGAGAGCTGCTGATTAAATCTTCCATCGAAGAGAGCTAATGAATCAATGTCATT	4380
Db	5354	GACGTGAGAGCTGCTGATTAAATCTTCCATCGAAGAGAGCTAATGAATCAATGTCATT	5413
Qy	4381	GCCAAAGATGTCAAAAGAAAGTTCTGATCTAGCAAAAGAAATTCAGGCTTTACTAAATGA	4440
Db	5414	GCCAAAGATGTCAAAAGAAAGTTCTGATCTAGCAAAAGAAATTCAGGCTTTACTAAATGA	5473
Qy	4441	AAAGTAAAAAATATTTTGAAGATCGTGTATTAATACGACTGGAATATGTTATTATACG	4500
Db	5474	AAAGTAAAAAATATTTTGAAGATCGTGTATTAATACGACTGGAATATGTTATTATACG	5533
Qy	4501	AAGGAACAATGAAAAAAGCAAAAGAAAAAGAGAGAGCGCTCATTTGTAATGCTGCTTTA	4560
Db	5534	AAGGAACAATGAAAAAAGCAAAAGAAAAAGAGAGAGCGCTCATTTGTAATGCTGCTTTA	5593
Qy	4561	TCGGTTCGTGAACGGATTAATCCGCTGAGAGTAGCTATTGCAATCAATCTGTTAAA	4620
Db	5594	TCGGTTCGTGAACGGATTAATCCGCTGAGAGTAGCTATTGCAATCAATCTGTTAAA	5653
Qy	4621	AATTAATTTAAGCAGAAATGAGTGAAGCAATTAAGAAAGCCGAGAGAGATAAATTCAT	4680
Db	5654	AATTAATTTAAGCAGAAATGAGTGAAGCAATTAAGAAAGCCGAGAGAGATAAATTCAT	5713
Qy	4681	GCGAAAACATGTAAATGTGGAAGCAAAATCATCTACTGTGTTGTGAATGCGGCTTCGGA	4740
Db	5714	GCGAAAACATGTAAATGTGGAAGCAAAATCATCTACTGTGTTGTGAATGCGGCTTCGGA	5773
Qy	4741	CTTGCTATCAGCAAAAGATGCTTTTCAAGAAATGGGATCTGAGACATGCAAGACTTATCA	4800
Db	5774	CTTGCTATCAGCAAAAGATGCTTTTCAAGAAATGGGATCTGAGACATGCAAGACTTATCA	5833
Qy	4801	AATGACACGATTTGCAAAAGTGTGATTAAGAAAGAAATTTCTGCTGATTCCTTAAATGTGAC	4860
Db	5834	AATGACACGATTTGCAAAAGTGTGATTAAGAAAGAAATTTCTGCTGATTCCTTAAATGTGAC	5893
Qy	4861	GCAAAATTAATTCATCTTGCGGGTGAATGTGCGGGAACATTCGCGGTTCTCTTCTACG	4920
Db	5894	GCAAAATTAATTCATCTTGCGGGTGAATGTGCGGGAACATTCGCGGTTCTCTTCTACG	5953
Qy	4921	GCGGTAGAGCTGCTTTTGCAGATAATCTCTCAATAAATAAATCTGCTTTGATTACA	4980
Db	5954	GCGGTAGAGCTGCTTTTGCAGATAATCTCTCAATAAATAAATCTGCTTTGATTACA	6013
Qy	4981	GGAACGAAGGTAAATCTTTTATGTGAAAGAAATCAAAAGTCAATGTATCAAGCTTTGAT	5040
Db	6014	GGAACGAAGGTAAATCTTTTATGTGAAAGAAATCAAAAGTCAATGTATCAAGCTTTGAT	6073

QY	5041	GATTCTCATATTACAAACGTTTCTGCTCGAGAGCGCTCCGACGTAATTAAGCAGCGCTGGATC	5100
Db	6074	GATTCCTCATATTACAAACGTTTCTGCTCGAGAGCGCTCCGACGTAATTAAGCAGCGCTGGATC	6133
QY	5101	GGAGGAAATGGATCTGCACATCGGCTTTCGATGTAACCGAAGCTTATAGTAGATCTC	5160
Db	6134	GGAGGAAATGGATCTGCATCTGTCATCGTGGTTCTGATGTAACCGAAGCTTATAGTAGATCTC	6193
QY	5161	GAGTTTGAAGGAGTAAAGTTCTTTCAATGTAGATGCAAAAGATCAAAAACATTAATACA	5220
Db	6194	GAGTTTGAAGGAGTAAAGTTCTTTCAATGTAGATGCAAAAGATCAAAAACATTAATACA	6253
QY	5221	ATTGCGGAAATGCAAAATGAGGAAAAACCGCTGGAAGTTGAGCAACAAGTTGCTATACA	5280
Db	6254	ATTGCGGAAATGCAAAATGAGGAAAAACCGCTGGAAGTTGAGCAACAAGTTGCTATACA	6313
QY	5281	AAATATGGAAACAATCACTGATTAATGCTATTGTAAAAAAAGTAAATTTACAAACGGCGAAT	5340
Db	6314	AAATATGGAAACAATCACTGATTAATGCTATTGTAAAAAAAGTAAATTTACAAACGGCGAAT	6373
QY	5341	GATCAAGATAGAAAAAATATCAATATGTACATGCAAAAGATTATACATATGCCAATATCATTA	5400
Db	6374	GATCAAGATAGAAAAAATATCAATATGTACATGCAAAAGATTATACATATGCCAATATCATTA	6433
QY	5401	GCAGTCGAGTTGGAGAGAGCAAAACGACCTCTGTGCAGAGAGCTTTCGACAGTACTAC	5460
Db	6434	GCAGTCGAGTTGGAGAGAGCAAAACGACCTCTGTGCAGAGAGCTTTCGACAGTACTAC	6493
QY	5461	TTGGAATTAAGACAGTTTCTTCTCAGTGTATCAACATATGACAAAGATTTAGAGGAA	5520
Db	6494	TTGGAATTAAGACAGTTTCTTCTCAGTGTATCAACATATGACAAAGATTTAGAGGAA	6553
QY	5521	GAAAAATATGAGAAATTAAGGAAAAAGCAAAATGTTATGTTCTAGCTGAAAAATACAGTCAA	5580
Db	6554	GAAAAATATGAGAAATTAAGGAAAAAGCAAAATGTTATGTTCTAGCTGAAAAATACAGTCAA	6613
QY	5581	GTGTCACAAATGCGACAGTCTTTCCGAGCAAGTGAACAAGCTGCAGTATGGAGCTGGA	5640
Db	6614	GTGTCACAAATGCGACAGTCTTTCCGAGCAAGTGAACAAGCTGCAGTATGGAGCTGGA	6673
QY	5641	GTAGCAGTTAATTAATTAACAACAATACTTCTGCACATTAATAAAAAATGATACCAAAAT	5700
Db	6674	GTAGCAGTTAATTAATTAACAACAATACTTCTGCACATTAATAAAAAATGATACCAAAAT	6733
QY	5701	GTAGCAAAATGCTTGGTGTAAAGCAAAATCTCATTCATTAATAACAATTTGGAATTGGA	5760
Db	6734	GTAGCAAAATGCTTGGTGTAAAGCAAAATCTCATTCATTAATAACAATTTGGAATTGGA	6793
QY	5761	GCTGAGTTGAGCTGAGAGAGCTGAGATGACAGGTTCTGTAGCAGTAAATTAAGATTGTA	5820
Db	6794	GCTGAGTTGAGCTGAGAGAGCTGAGATGACAGGTTCTGTAGCAGTAAATTAAGATTGTA	6853
QY	5821	AATTAATACGATAGCAGAAATTAATATCATGCAAAAATCACTGCGAAGGAAATGTCGGAGTT	5880
Db	6854	AATTAATACGATAGCAGAAATTAATATCATGCAAAAATCACTGCGAAGGAAATGTCGGAGTT	6913
QY	5881	ATTACAGAGTCGATAGCGGTAATGCTAAATATGACAGGAACAGTGTCTGAGTGGACCGGT	5940
Db	6914	ATTACAGAGTCGATAGCGGTAATGCTAAATATGACAGGAACAGTGTCTGAGTGGACCGGT	6973
QY	5941	GCAGCAATAGAGCCTCAACCAAGTGTGATGAAATTACAGAGATCTCAAAAAGCATATGTA	6000
Db	6974	GCAGCAATAGAGCCTCAACCAAGTGTGATGAAATTACAGAGATCTCAAAAAGCATATGTA	7033
QY	6001	AAAGATTCTACAGTGAATGCTTAAGAGAAACAGATGATTAATTAATCTACCAAGGGGAA	6060
Db	7034	AAAGATTCTACAGTGAATGCTTAAGAGAAACAGATGATTAATTAATCTACCAAGGGGAA	7093
QY	6061	GTAGATTAAGTGTGATTAAGTATTCAAAATCTTAATTAATTAAGAGAGCTTATACAA	6120
Db	7094	GTAGATTAAGTGTGATTAAGTATTCAAAATCTTAATTAATTAAGAGAGCTTATACAA	7153
QY	6121	AAAAAATAAGTAAATTAATAAAAGAAATTTGTTACCAATAGTCACTACTCATATCTTTA	6180

Db 7154 AAAAAAAAAAATAAGATTAATAAAAAAAAAAGATTGTTACCAATAGTTCAGCTACCTCATCTTTA 7213
Qy 6181 AATCTTTATTTGCAAAATGCGCTGGTTTCAAGACAAGCGAGTGGCAGAACTGTAT 6240
Db 7214 AATCTTTATTTGCAAAATGCGCTGGTTTCAAGACAAGCGAGTGGCAGAACTGTAT 7273
Qy 6241 ATCAACAAGGTTTATGAGAAAACAGAAGCTCTTGAGAAAAATCTATAATTAAATGCAAA 6300
Db 7274 ATCAACAAGGTTTATGAGAAAACAGAAGCTCTTGAGAAAAATCTATAATTAAATGCAAA 7333
Qy 6301 CATTTATTTCTGTAATTCAGAGATTTACAGAAATTCATTCGAGTAGTAGGTTCTGTGCT 6360
Db 7334 CATTTATTTCTGTAATTCAGAGATTTACAGAAATTCATTCGAGTAGTAGGTTCTGTGCT 7393
Qy 6361 GTTGTGGAATGTAGAGTAGAGCTCTTCTGATACCAATATTAATAAAGAAATAC 6420
Db 7394 GTTGTGGAATGTAGAGTAGAGCTCTTCTGATACCAATATTAATAAAGAAATAC 7453
Qy 6421 AAGACAAGGTTGAAAAAATCAAAATGCTGATGAAGGTTTCGAGAAAGCTGAAT 6480
Db 7454 AAGACAAGGTTGAAAAAATCAAAATGCTGATGAAGGTTTCGAGAAAGCTGAAT 7513
Qy 6481 ACAGCAGATTCTAACAGAGAAATTTCTCTTTTGAAGTCGAGTCGACAGCGGGTA 6540
Db 7514 ACAGCAGATTCTAACAGAGAAATTTCTCTTTTGAAGTCGAGTCGAGTCGACAGCGGGTA 7573
Qy 6541 GGAGCCGAGTGGCAGAGACCGTTCCGTAATCAATTTGACAGAAAGCGGAAGTAGAT 6600
Db 7574 GGAGCCGAGTGGCAGAGACCGTTCCGTAATCAATTTGACAGAAAGCGGAAGTAGAT 7633
Qy 6601 GTGAGAAGAGCAAGATTGTTGTAAAAAAAGCTGAGATTACAGCAAAACGTTATAGTCT 6660
Db 7634 GTGAGAAGAGCAAGATTGTTGTAAAAAAAGCTGAGATTACAGCAAAACGTTATAGTCT 7693
Qy 6661 GTTGCAATTTGAAAAATGCGCAGTCGAGTAGCTGCAAAAGAGCTGGAATTTGACAGCA 6720
Db 7694 GTTGCAATTTGAAAAATGCGCAGTCGAGTAGCTGCAAAAGAGCTGGAATTTGACAGCA 7753
Qy 6721 GTGAGAGTTTACCAAGATGAAATCAAAACAGAGCAAGAGTGAATAATTTCTAAATTTAG 6780
Db 7754 GTGAGAGTTTACCAAGATGAAATCAAAACAGAGCAAGAGTGAATAATTTCTAAATTTAG 7813
Qy 6781 ACTCGAAAACAAGTTAGATGATTAAGCAGAAAAAGATTAATCAAGTAACTGGAATCGGT 6840
Db 7814 ACTCGAAAACAAGTTAGATGATTAAGCAGAAAAAGATTAATCAAGTAACTGGAATCGGT 7873
Qy 6841 TCAGCCGAGCTGGAATTTCTGACAGCCGAGATTCGAGAGTGTCTGTCAATTAATTT 6900
Db 7874 TCAGCCGAGCTGGAATTTCTGACAGCCGAGATTCGAGAGTGTCTGTCAATTAATTT 7933
Qy 6901 GCAAAATTAAGTAGAAGATGATGATCACTAGTACTCTTCACTCTTCAAGTGAAT 6960
Db 7934 GCAAAATTAAGTAGAAGATGATGATCACTAGTACTCTTCACTCTTCAAGTGAAT 7993
Qy 6961 GTTAAAGCTCTTAATTAATTTGAAATTCCTTGAAGCCGAGTGAAGAGCCGAGTCTT 7020
Db 7994 GTTAAAGCTCTTAATTAATTTGAAATTCCTTGAAGCCGAGTGAAGAGCCGAGTCTT 8053
Qy 7021 GCAGCAGTTTACCGAGTGTCTTCTGTAAACATTAATAGTTCTGTGATAGCTGAGTT 7080
Db 8054 GCAGCAGTTTACCGAGTGTCTTCTGTAAACATTAATAGTTCTGTGATAGCTGAGTT 8113
Qy 7081 CACAATTAATCTGATTTGATCTCCGTAAGAGAAAAAGTAAATGTAACGGCAAAAGAGAA 7140
Db 8114 CACAATTAATCTGATTTGATCTCCGTAAGAGAAAAAGTAAATGTAACGGCAAAAGAGAA 8173
Qy 7141 AAAAAATTTAAGCAAAACAGCAGCAAAATCGAGAAATCGAGAGCAGCAATCGAGCCAT 7200
Db 8174 AAAAAATTTAAGCAAAACAGCAGCAAAATCGAGAAATCGAGAGCAGCAATCGAGCCAT 8233
Qy 7201 GTCTTGTAAATTAATTTTGAACAGCTGTGAAGATGAAAAAATTTCTGAAGAAAGAG 7260

Db 8234 GTCTTGTAAATTAATTTTGAACAGCTGTGAAGATGAAAAAATTTCTGAAGAAAGAG 8293
Qy 7261 ACAGAAAGTTTAAAAAATTTTGAAGCAAGTTAAACAAGAACAGATTAATAAATGAT 7320
Db 8294 ACAGAAAGTTTAAAAAATTTTGAAGCAAGTTAAACAAGAACAGATTAATAAATGAT 8353
Qy 7321 GCTACGAAAAAATCTTAACAATCAGCAGATATTTCTACAGAAATCTCTGTAAACGG 7380
Db 8354 GCTACGAAAAAATCTTAACAATCAGCAGATATTTCTACAGAAATCTCTGTAAACGG 8413
Qy 7381 GATAGAGAGTACTCAGAGGAGAAAGAAATTAACCAATGAGAGACTCTGTATTTAT 7440
Db 8414 GATAGAGAGTACTCAGAGGAGAAAGAAATTAACCAATGAGAGACTCTGTATTTAT 8473
Qy 7441 GGAATAAATGTAGATTTATTAACAACAGAGCAAGAAATATCATCTTCTGATGCTTTG 7500
Db 8474 GGAATAAATGTAGATTTATTAACAACAGAGCAAGAAATATCATCTTCTGATGCTTTG 8533
Qy 7501 GGAATCGCAGGCTTTCCTCCGATCAGGAACAGTGGCAGTTACAAATATTTAAAGAAAT 7560
Db 8534 GGAATCGCAGGCTTTCCTCCGATCAGGAACAGTGGCAGTTACAAATATTTAAAGAAAT 8593
Qy 7561 TCCGAGATTACTGTGAAAAATTTCTTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGATCG 7620
Db 8594 TCCGAGATTACTGTGAAAAATTTCTTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGATCG 8653
Qy 7621 GATATTAACAGAAATGTGCTTTAAACAGCATATCAAGGTCCTGTAGAGCATTTGGAAATA 7680
Db 8654 GATATTAACAGAAATGTGCTTTAAACAGCATATCAAGGTCCTGTAGAGCATTTGGAAATA 8713
Qy 7681 GGAGCTGCTATGAGAAATTAATTTCTAATGAAAGATCAAAATATCAGTATTAATAATCT 7740
Db 8714 GGAGCTGCTATGAGAAATTAATTTCTAATGAAAGATCAAAATATCAGTATTAATAATCT 8773
Qy 7741 AAGCTATTAAGAAAAAATATTTGATGTTATTTGTAAGTAAATGCGAAATTTGAGCCGAA 7800
Db 8774 AAGCTATTAAGAAAAAATATTTGATGTTATTTGTAAGTAAATGCGAAATTTGAGCCGAA 8833
Qy 7801 GCAAAAAGATTAACCGTAGAGAGCGGTAGCTCCGAGGACATATCTCAAAAGCAAAAGAT 7860
Db 8834 GCAAAAAGATTAACCGTAGAGAGCGGTAGCTCCGAGGACATATCTCAAAAGCAAAAGAT 8893
Qy 7861 GAAATGAATTCAGAGGTTGAAATTTGAAGAGATATTTCAATGAAGAAATAGAGTAACT 7920
Db 8894 GAAATGAATTCAGAGGTTGAAATTTGAAGAGATATTTCAATGAAGAAATAGAGTAACT 8953
Qy 7921 AGCCCTTCTAAAGAAATTTGGAAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGTG 7980
Db 8954 AGCCCTTCTAAAGAAATTTGGAAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGTG 9013
Qy 7981 ACTGCTGAATCTCAAGAGGCTTCTGTAGAGCAGTAGAGGAGGAGAAATTAATTCGAA 8040
Db 9014 ACTGCTGAATCTCAAGAGGCTTCTGTAGAGCAGTAGAGGAGGAGAAATTAATTCGAA 9073
Qy 8041 GCAAAAAGTCCGAGAGCTCTTAATTTGAAGTAACTAGTCAAAATCCGAGAAAGATTTTT 8100
Db 9074 GCAAAAAGTCCGAGAGCTCTTAATTTGAAGTAACTAGTCAAAATCCGAGAAAGATTTTT 9133
Qy 8101 CATGCAAGTATGTGAATATGGAAGCAACATTAATTAAGAAAGTAACGAGCTTTCTAAA 8160
Db 9134 CATGCAAGTATGTGAATATGGAAGCAACATTAATTAAGAAAGTAACGAGCTTTCTAAA 9193
Qy 8161 GCAGTAACAGGTTCTGTATTTGAGAGAGGTTGAGTGCACCAAGGAGAAAGTAACTGCTGCA 8220
Db 9194 GCAGTAACAGGTTCTGTATTTGAGAGAGGTTGAGTGCACCAAGGAGAAAGTAACTGCTGCA 9253
Qy 8221 GGTAAAACTATGTAGAAAGTTGAGAAAGAAATTTGTTCAAGAACAAATCGATTGAATGCA 8280
Db 9254 GGTAAAACTATGTAGAAAGTTGAGAAAGAAATTTGTTCAAGAACAAATCGATTGAATGCA 9313
Qy 8281 ATTTCTAAAGTAAGAGTTTGTGATGAAGATTAAGTAACTGCTAAATCTTCTGTAGTATCA 8340
Db 9314 ATTTCTAAAGTAAGAGTTTGTGATGAAGATTAAGTAACTGCTAAATCTTCTGTAGTATCA 9373

QY 8341 GGAATGAGAGAGATTCCCGAGCAGAGTGAATCTTCTACAGCACAAGTAATPCT 8400
| | | | |
DB 9374 GGAATGAGAGAGATTCCCGAGCAGAGTGAATCTTCTACAGCACAAGTAATPCT 9433
| | | | |
QY 8401 GAAATCCGATTCGTTTACGAAGCAGATTAATGAAATTAATGATTAACAAGAAAAATAT 8460
| | | | |
DB 9434 GAAATCCGATTCGTTTACGAAGCAGATTAATGAAATTAATGATTAACAAGAAAAATAT 9493
| | | | |
QY 8461 ATTTGAGAGTCAATGCTCTTCTTAAATGATCAAGAAATGAAGCAATATAGAAATCT 8520
| | | | |
DB 9494 ATTTGAGAGTCAATGCTCTTCTTAAATGATCAAGAAATGAAGCAATATAGAAATCT 9553
| | | | |
QY 8521 TTAGCGTAGCCGCTGTCATGACACAGCAACAAAGCAATTAAGATCAAGCAAG 8580
| | | | |
DB 9554 TTAGCGTAGCCGCTGTCATGACACAGCAACAAAGCAATTAAGATCAAGCAAG 9613
| | | | |
QY 8581 TTAACTTCTACAACTGTAATGAGAGAAACGATCTCAACTTCGTCAAAAGCTTTGGCT 8640
| | | | |
DB 9614 TTAACTTCTACAACTGTAATGAGAGAAACGATCTCAACTTCGTCAAAAGCTTTGGCT 9673
| | | | |
QY 8641 AAAAATGAAATTTATGAAATGTAAGGAACTGAGAGACCTTAGTCGAGCGGAAACA 8700
| | | | |
DB 9674 AAAAATGAAATTTATGAAATGTAAGGAACTGAGAGACCTTAGTCGAGCGGAAACA 9733
| | | | |
QY 8701 GCAGCGTTGAAATTTATACAAAGTACTACAGAGCAATGCTTGACAGAAATTTGGAA 8760
| | | | |
DB 9734 GCAGCGTTGAAATTTATACAAAGTACTACAGAGCAATGCTTGACAGAAATTTGGAA 9793
| | | | |
QY 8761 ATTGAGATTAATTTGAAACGATTTGCAAGAGTAATAGATTTGAAAGTCAAGGAGAC 8820
| | | | |
DB 9794 ATTGAGATTAATTTGAAACGATTTGCAAGAGTAATAGATTTGAAAGTCAAGGAGAC 9853
| | | | |
QY 8821 GGAACCAAGAGAGCTTGTCCGAAAGATGATTTCTGTGAAATTAACAATTTCAAGG 8880
| | | | |
DB 9854 GGAACCAAGAGAGCTTGTCCGAAAGATGATTTCTGTGAAATTAACAATTTCAAGG 9913
| | | | |
QY 8881 GAAACCAAAATCATCATTTGAAAGATTAAGCAGAAATTTGTAACCGGAAAGTGAATGTA 8940
| | | | |
DB 9914 GAAACCAAAATCATCATTTGAAAGATTAAGCAGAAATTTGTAACCGGAAAGTGAATGTA 9973
| | | | |
QY 8941 GATGCTTTGAAATGAACTTGATGATGATCTACAGAGAAAAAGTGGCTAATGCTGAAT 9000
| | | | |
DB 9974 GATGCTTTGAAATGAACTTGATGATGATCTACAGAGAAAAAGTGGCTAATGCTGAAT 10033
| | | | |
QY 9001 GGTATTGGAATGTTGATGATTAATATGATGATTAAGAAAAATGTGAAGCCAAAAATCGA 9060
| | | | |
DB 10034 GGTATTGGAATGTTGATGATTAATATGATGATTAAGAAAAATGTGAAGCCAAAAATCGA 10093
| | | | |
QY 9061 AGACATGCTATTGTGAAAACTACGTAAGAAAAACAAGATTAACAACATTTTCAAGAGCAAA 9120
| | | | |
DB 10094 AGACATGCTATTGTGAAAACTACGTAAGAAAAACAAGATTAACAACATTTTCAAGAGCAAA 10153
| | | | |
QY 9121 GTAATATTTCTTGAAAAAGAGACGCTGACGCTGCAATATCGAATGTAACATTT 9180
| | | | |
DB 10154 GTAATATTTCTTGAAAAAGAGACGCTGACGCTGCAATATCGAATGTAACATTT 10213
| | | | |
QY 9181 TCCATATGATGATTAATTAATTTGGCAAGCAGATGATGATCTTCAATTAATPACC 9240
| | | | |
DB 10214 TCCATATGATGATTAATTAATTTGGCAAGCAGATGATGATCTTCAATTAATPACC 10273
| | | | |
QY 9241 AAAAATTCAAAAAATTAATTAATTTAGCATCAAGTATGAAATGTAATGTTTCAT 9300
| | | | |
DB 10274 AAAAATTCAAAAAATTAATTAATTTAGCATCAAGTATGAAATGTAATGTTTCAT 10333
| | | | |
QY 9301 GGGGTGCTGAGCAAGAGTGACAGAGCCAAAGCAGATTAGTGAAGATCAAAAT 9360
| | | | |
DB 10334 GGGGTGCTGAGCAAGAGTGACAGAGCCAAAGCAGATTAGTGAAGATCAAAAT 10393
| | | | |
QY 9361 AATGAACTTAATATGTTGATTTAGCAGAGAAAAATTAACAAGAGGAAACATCAATGTA 9420
| | | | |
DB 10394 AATGAACTTAATATGTTGATTTAGCAGAGAAAAATTAACAAGAGGAAACATCAATGTA 10453
| | | | |

QY 9421 TATCCGGATATGATTAATTAATTAATTAATGATTAAGACAAATTTCAAGCTATTCCGAT 9480
| | | | |
DB 10454 TATCCGGATATGATTAATTAATTAATTAATGATTAAGACAAATTTCAAGCTATTCCGAT 10513
| | | | |
QY 9481 GCCAAAGTCAATGCTGAGCTGCTTGGCAATGCACTGCACTATTGAAAAAATGAAGTAAA 9540
| | | | |
DB 10514 GCCAAAGTCAATGCTGAGCTGCTTGGCAATGCACTGCACTATTGAAAAAATGAAGTAAA 10573
| | | | |
QY 9541 TTTTAATATGATCCGAGAAATTTAAAAATTAATTCGCAAGATTGGAAGGGAAGCTAAT 9600
| | | | |
DB 10574 TTTTAATATGATCCGAGAAATTTAAAAATTAATTCGCAAGATTGGAAGGGAAGCTAAT 10633
| | | | |
QY 9601 AAAAAACGTGATGATCTTAATCAGTACGATGATATACGATTAATATATACATGCGAT 9660
| | | | |
DB 10634 AAAAAACGTGATGATCTTAATCAGTACGATGATATACGATTAATATATACATGCGAT 10693
| | | | |
QY 9661 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAGAGAGGAAAAAGGAAA 9720
| | | | |
DB 10694 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAGAGAGGAAAAAGGAAA 10753
| | | | |
QY 9721 AATGCA 9726
| | | | |
DB 10754 AATGCA 10759
| | | | |
RESULT 4
US-10-647-057-15
; Sequence 15, Application US/10647057
; Publication No. US2004004787A1
; GENERAL INFORMATION:
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION
; APPLICANT: STEWART, GEORGE
; APPLICANT: NARAYANAN, SANJEEV
; APPLICANT: CHENGARPA, M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP.
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: PCT-30962
; CURRENT APPLICATION NUMBER: US/10/647.057
; CURRENT FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 11130
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-10-647-057-15
Query Match 100.0%; Score 9726; DB 7; Length 11130;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGGATCAAAAAATTAACGTTGAGAGACAAAGAGAGATATCAGATTCTAAAAA 60
| | | | |
DB 1034 ATGAGCGGATCAAAAAATTAACGTTGAGAGACAAAGAGAGATATCAGATTCTAAAAA 1093
| | | | |
QY 61 GTTTTAATGATTTGGATGTTGATTAACACTATGACGCTGAGGCTTAATGATCAATC 120
| | | | |
DB 1094 GTTTTAATGATTTGGATGTTGATTAACACTATGACGCTGAGGCTTAATGATCAATC 1153
| | | | |
QY 121 ACCGCGATGAGAAATTTTGGAACAAAAATGAAAAAAGGATTAATGATGACTTACT 180
| | | | |
DB 1154 ACCGCGATGAGAAATTTTGGAACAAAAATGAAAAAAGGATTAATGATGACTTACT 1213
| | | | |
QY 181 ACAACCAAGATTCAGAGGAGAACGCTTTTAACAGTTTAATGATTTGCTTTAACAGAA 240
| | | | |
DB 1214 ACAACCAAGATTCAGAGGAGAACGCTTTTAACAGTTTAATGATTTGCTTTAACAGAA 1273
| | | | |
QY 241 AATTAATATGCAATCTATATTTTGGGAAAAAGATATGATCGGGGTAAATATCTTTT 300
| | | | |
DB 1274 AATTAATATGCAATCTATATTTTGGGAAAAAGATATGATCGGGGTAAATATCTTTT 1333
| | | | |
QY 301 AACTTTGCAATGAAAAATTAAGTATGAGATTAACAAGGAATTCAGAAAAATAA 360
| | | | |

Db 1334 AACCTTGTCAATGGAATAATGAGTAGATGGATTATCAACGGAATTCGAGAAATATAA 1393
Qy 1361 ATTGGAGGAATTTATATTTCTTAAAGCTCGGAAGGATGGCAGTAGGAATAATGAGATT 420
Db 1394 ATTGGAGGAATTTATATTTCTTAAAGCTCGGAAGGATGGCAGTAGGAATAATGAGATT 1453
Qy 421 ATCAAATGCTGCTCTTTTCATTCATTTATTTCCAAACCAAGATGATTTTAAAGAGCTTTG 480
Db 1454 ATCAATGCTGCTCTTTTCATTCATTTATTTCCAAACCAAGATGATTTTAAAGAGCTTTG 1513
Qy 481 GAAAGAACCAACATGCTAAAGTTTAAATGGAATCATTCAGTGAATGGAATAAGTAAA 540
Db 1514 GAAAGAACCAACATGCTAAAGTTTAAATGGAATCATTCAGTGAATGGAATAAGTAAA 1573
Qy 541 ATTCATTTGAATCCGAATGGAAGCATTAAGTAGAAGGAATAATCAATGCTGTGAAGC 600
Db 1574 ATTCATTTGAATCCGAATGGAAGCATTAAGTAGAAGGAATAATCAATGCTGTGAAGC 1633
Qy 601 ATCGGTTTATATGCGCGGATATTAATGGAAGAATCTCAATATCTAAAGAAGAAAT 660
Db 1634 ATCGGTTTATATGCGCGGATATTAATGGAAGAATCTCAATATCTAAAGAAGAAAT 1693
Qy 661 ACAGATTTTAAATTTTATGCTAATATTAATGATCGAATTAATTTGCTGTGAACGAGAT 720
Db 1694 ACAGATTTTAAATTTTATGCTAATATTAATGATCGAATTAATTTGCTGTGAACGAGAT 1753
Qy 721 TTAAGAGCTACCAAGCAAAATCTGAGATATTAATTTCTTGAGCTCACATGATCTCT 780
Db 1754 TTAAGAGCTACCAAGCAAAATCTGAGATATTAATTTCTTGAGCTCACATGATCTCT 1813
Qy 781 CAAAAAGCTATGGAAAAAATTCATCTGTGGAAGAGAAATAGAAATATGTAAAGAA 840
Db 1814 CAAAAAGCTATGGAAAAAATTCATCTGTGGAAGAGAAATAGAAATATGTAAAGAA 1873
Qy 841 AATACCAAGCAAAATATGGAATCTGATGTTTGAAGAGAGATGGAATATTAATAAT 900
Db 1874 AATACCAAGCAAAATATGGAATCTGATGTTTGAAGAGAGATGGAATATTAATAAT 1933
Qy 901 AGTGGAAAGCTACAAATGGGAGATTTTAAGAAAGAAAGGGAAGAAAGAACTTATAC 960
Db 1934 AGTGGAAAGCTACAAATGGGAGATTTTAAGAAAGAAAGGGAAGAAAGAACTTATAC 1993
Qy 961 ACTCCTTTAAGTTTATCAGATGTGAAGCTTCCTGAAGATTAATTAAGAAAGATCATA 1020
Db 1994 ACTCCTTTAAGTTTATCAGATGTGAAGCTTCCTGAAGATTAATTAAGAAAGATCATA 2053
Qy 1021 GGAAGAAATGTTGACATTAAGCTGAAGCAAGAAATTTCTATGATGCAACTTATGTA 1080
Db 2054 GGAAGAAATGTTGACATTAAGCTGAAGCAAGAAATTTCTATGATGCAACTTATGTA 2113
Qy 1081 AAGCTTGCAAGACACTTTTAGCTTTGTTACAGGTTCAATTTCTCTATCAATTTAAT 1140
Db 2114 AAGCTTGCAAGACACTTTTAGCTTTGTTACAGGTTCAATTTCTCTATCAATTTAAT 2173
Qy 1141 GGAATTTTAAAGTTTATGCAAGTAAGTCCAGTGTCTTATGGAAGAAAGATCCAAAGT 1200
Db 2174 GGAATTTTAAAGTTTATGCAAGTAAGTCCAGTGTCTTATGGAAGAAAGATCCAAAGT 2233
Qy 1201 GAAAGCAAGAAAGAAAGCAATATTCATTTCAAGTGAAGTAAGCAACTATGGA 1260
Db 2234 GAAAGCAAGAAAGAAAGCAATATTCATTTCAAGTGAAGTAAGCAACTATGGA 2293
Qy 1261 GAGGCTACTCTTCCATTAATAATTTCAATTTATTTGAGAGAAAGCCAAATGGAATCT 1320
Db 2294 GAGGCTACTCTTCCATTAATAATTTCAATTTATTTGAGAGAAAGCCAAATGGAATCT 2353
Qy 1321 CTCAGTATCGAGCGGATATATTTCTGCAAAAAAGTAATTCATATGTAATCTTGAAGA 1380
Db 2354 CTCAGTATCGAGCGGATATATTTCTGCAAAAAAGTAATTCATATGTAATCTTGAAGA 2413
Qy 1381 GAAATTAATGAAAGGAAAGACAGATATTAATCTGAAATATCTATGATGCT 1440

Db 2414 GAAATTAATGAAAGGAAAGACAGATATTAATCTCAAAATCTGAAAAATCTATGATGCT 2473
Qy 1441 TCTGTTTCTGTGGAAGCATGAGATTCATTAATAAGATCTTTAGATATGATGACG 1500
Db 2474 TCTGTTTCTGTGGAAGCATGAGATTCATTAATAAGATCTTTAGATATGATGACG 2533
Qy 1501 GAAAGAAATAATCTTCCGTCAGATTTGCTAAGAGCAAAAGTAAGATCAAGAAACG 1560
Db 2534 GAAAGAAATAATCTTCCGTCAGATTTGCTAAGAGCAAAAGTAAGATCAAGAAACG 2593
Qy 1561 GATGATGTAATGTAAGCAAGTGAAGCAATTAATTCATTTGAGTGTCTGTAAAGTGA 1620
Db 2594 GATGATGTAATGTAAGCAAGTGAAGCAATTAATTCATTTGAGTGTCTGTAAAGTGA 2653
Qy 1621 TTGGGGGATAGTGTAAATGGGGTGTGGCTGCAATATTTTCTAATCTAATGCTTCTCC 1680
Db 2654 TTGGGGGATAGTGTAAATGGGGTGTGGCTGCAATATTTTCTAATCTAATGCTTCTCC 2713
Qy 1681 CGTATGATGTAATGATATCTACATGCAAGAGGACCTAAATGTGAGGCTCATTAAC 1740
Db 2714 CGTATGATGTAATGATATCTACATGCAAGAGGACCTAAATGTGAGGCTCATTAAC 2773
Qy 1741 ATTAATAATAATAGTGTCTGCAACAGATCTGATTTGGGAACCTTCCAGTTTATGAAT 1800
Db 2774 ATTAATAATAATAGTGTCTGCAACAGATCTGATTTGGGAACCTTCCAGTTTATGAAT 2853
Qy 1801 GATCAGCTTTATGAATCAGGTCATCTAAATCAATTTTATGATGCAATTAACAGCGGTT 1860
Db 2854 GATCAGCTTTATGAATCAGGTCATCTAAATCAATTTTATGATGCAATTAACAGCGGTT 2893
Qy 1861 GGAGAGACAGTGTCAATGAGAAATTAAGATTAAGCTTAACGAATTAATTAATGTCGCT 1920
Db 2894 GGAGAGACAGTGTCAATGAGAAATTAAGATTAAGCTTAACGAATTAATTAATGTCGCT 2953
Qy 1921 GTGTCTGCAACCATAGCAAAATCATATTAATTTCTGCTCTGTGCAATGAGAGATGGA 1980
Db 2954 GTGTCTGCAACCATAGCAAAATCATATTAATTTCTGCTCTGTGCAATGAGAGATGGA 3013
Qy 1981 AGACTTTCTTCAAGAGTGAAGGAGATTAATTAAGGAGCATTAATTAAGGCTCAAAATCTT 2040
Db 3014 AGACTTTCTTCAAGAGTGAAGGAGATTAATTAAGGAGCATTAATTAAGGCTCAAAATCTT 3073
Qy 2041 CGAGCAGTACGTCAGTGAAGTGTGCTGTACGAAAGGAAGAAAGAAAGAACTTAT 2100
Db 3074 CGAGCAGTACGTCAGTGAAGTGTGCTGTACGAAAGGAAGAAAGAAAGAAAGAACTTAT 3133
Qy 2101 GGAATGAGAGAGTTTATATGAAATCATATTAATTAATTTCTGTGCAATTTGCCGAT 2160
Db 3134 GGAATGAGAGAGTTTATATGAAATCATATTAATTAATTTCTGTGCAATTTGCCGAT 3193
Qy 2161 CATGCTGAATTTGATGCGAAGAAATTAATTAAGCAAGTGAATTAATTAATTAAT 2220
Db 3194 CATGCTGAATTTGATGCGAAGAAATTAATTAAGCAAGTGAATTAATTAATTAAT 3253
Qy 2221 AAAAAATCTTCAAAAAATGCAAGTCTGTTATTAATTAATTAATTAATTAATTAAGAGCT 2280
Db 3254 AAAAAATCTTCAAAAAATGCAAGTCTGTTATTAATTAATTAATTAATTAATTAAGAGCT 3313
Qy 2281 TTGGAAGAAAGAAAGAAAGCTCCAGAAATTAATTAATTAATTAATTAATTAATTAAG 2340
Db 3314 TTGGAAGAAAGAAAGAAAGCTCCAGAAATTAATTAATTAATTAATTAATTAATTAAG 3373
Qy 2341 TTATTTGAATGCAATTTTCAAGAAATTTGATGGAAGAAAGCGAGCTTTTACTAAATGATGA 2400
Db 3374 TTATTTGAATGCAATTTTCAAGAAATTTGATGGAAGAAAGCGAGCTTTTACTAAATGATGA 3433
Qy 2401 AGAATGACAAATTAATTTCTCCGATGGAATCTTCAAAACAGGAATGCTATGAAATTTGCA 2460
Db 3434 AGAATGACAAATTAATTTCTCCGATGGAATCTTCAAAACAGGAATGCTATGAAATTTGCA 3493
Qy 2461 AACTATGTTCAAGGAGAAATGAATAATTAAGAGAAATAATTAAGAAAGATTTAAAGCT 2520
Db 3494 AACTATGTTCAAGGAGAAATGAATAATTAAGAGAAATAATTAAGAAAGATTTAAAGCT 3553

QY 2521 TTTTCAGAAAGATTGAGTGAAGTCAATTAAGAACTTTGAATTTTACAGAGTAGAAT 2580
 Db 3554 TTTTCAGAAAGATTGAGTGAAGTCAATTAAGAACTTTGAATTTTACAGAGTAGAAT 3613
 QY 2581 TATGAAATTTTCACTTTTACCTTCCGGAGTAAATGAGAAAGAGTCTTCT 2640
 Db 3614 TATGAAATTTTCACTTTTACCTTCCGGAGTAAATGAGAAAGAGTCTTCT 3673
 QY 2641 GTGGAGAGAGCTGTTCGTGGGTAGACAGAGAAATTAAGCAAGTATCCGTTGAAAA 2700
 Db 3674 GTGGAGAGAGCTGTTCGTGGGTAGACAGAGAAATTAAGCAAGTATCCGTTGAAAA 3733
 QY 2701 GAGACTAACTTGTCTGCAAAAAAAGATTAAATATAAAGCTATCAATTAAGCAGAAACA 2760
 Db 3734 GAGACTAACTTGTCTGCAAAAAAAGATTAAATATAAAGCTATCAATTAAGCAGAAACA 3793
 QY 2761 GTGAATTTAGTTGGAAAAATTTGGCACTTGGAGAACAGTACATCCGGAAAGTCCAGTCGGA 2820
 Db 3794 GTGAATTTAGTTGGAAAAATTTGGCACTTGGAGAACAGTACATCCGGAAAGTCCAGTCGGA 3853
 QY 2821 GGAAGATTAAATGTTCAAAGATCGAAAAATTCAGCTATCGTAGAAGCTTAAAGAAAAAGCT 2880
 Db 3854 GGAAGATTAAATGTTCAAAGATCGAAAAATTCAGCTATCGTAGAAGCTTAAAGAAAAAGCT 3913
 QY 2881 GAATTAATCAGAGAAAAATTAATATGCAATGCAATGCAAGCACTTTTCAATGTCGGAGA 2940
 Db 3914 GAATTAATCAGAGAAAAATTAATATGCAATGCAATGCAAGCACTTTTCAATGTCGGAGA 3973
 QY 2941 TCTTTTAAATGCTGCTCAGGTGGGAATGCAATCAATGGAATGGGAAGTTATGTGAAGT 3000
 Db 3974 TCTTTTAAATGCTGCTCAGGTGGGAATGCAATCAATGGAATGGGAAGTTATGTGAAGT 4033
 QY 3001 ATCAGTAAAGGCAAGATTTCATTAATGACGAAGCAATTTGAAAAAGCTTATAAAAAATT 3060
 Db 4034 ATCAGTAAAGGCAAGATTTCATTAATGACGAAGCAATTTGAAAAAGCTTATAAAAAATT 4093
 QY 3061 GCTTTAAACAGTAAGAAATGATCTTCTGTGGAAATGCTCCGGTTACCGGGAATCGGA 3120
 Db 4094 GCTTTAAACAGTAAGAAATGATCTTCTGTGGAAATGCTCCGGTTACCGGGAATCGGA 4153
 QY 3121 AGCAAAAAATGCGCGGCTCGGGGTGCTGTGGCGTAAATGATTAAGATTAATTTCAACAA 3180
 Db 4154 AGCAAAAAATGCGCGGCTCGGGGTGCTGTGGCGTAAATGATTAAGATTAATTTCAACAA 4213
 QY 3181 GCTTCATTTGAAGATTAATGACGAAGCAAGTAAATATGATTAAGATTAAGATTAAG 3240
 Db 4214 GCTTCATTTGAAGATTAATGACGAAGCAAGTAAATATGATTAAGATTAAGATTAAG 4273
 QY 3241 GTTAACAGTAAGTCCGGAATCTTTAGAGTAAATGCAAAAAAGCCGGAACAATCAACAGT 3300
 Db 4274 GTTAACAGTAAGTCCGGAATCTTTAGAGTAAATGCAAAAAAGCCGGAACAATCAACAGT 4333
 QY 3301 ATTTCTGTTCGGGAGAAATTAATTAAGTTGGAAGTAAACCGAGTAAAGAAAAAGCGAA 3360
 Db 4334 ATTTCTGTTCGGGAGAAATTAATTAAGTTGGAAGTAAACCGAGTAAAGAAAAAGCGAA 4393
 QY 3361 TCAGAAAGAAAGACCAAGGGAATTTTGGCAAAATCGGAAACAAGTGAATCTGTGAAA 3420
 Db 4394 TCAGAAAGAAAGACCAAGGGAATTTTGGCAAAATCGGAAACAAGTGAATCTGTGAAA 4453
 QY 3421 AATTAATAATTAAGGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3480
 Db 4454 AATTAATAATTAAGGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 4513
 QY 3481 GGAATTAATAATTAAGGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3540
 Db 4514 GGAATTAATAATTAAGGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 4573
 QY 3541 TCTTTCAAGTTGGAGCTTCTGGAAGTGTTCCTTTCAATTAATTAATTAATTAATTAATTA 3600
 Db 4574 TCTTTCAAGTTGGAGCTTCTGGAAGTGTTCCTTTCAATTAATTAATTAATTAATTAATTA 4633

QY 3601 GCTGCTGAGATGAGATTAAGATTAATTTGAAGGAGCAAAATTAAGTAGAGTGAAT 3660
 Db 4634 GCTGCTGAGATGAGATTAAGATTAATTTGAAGGAGCAAAATTAAGTAGAGTGAAT 4693
 QY 3661 TCTTCTGATTTCACTTTTGTGTGAGCAATGGGCGGATCTGTCGACTTCACTGGAATCAT 3720
 Db 4694 TCTTCTGATTTCACTTTTGTGTGAGCAATGGGCGGATCTGTCGACTTCACTGGAATCAT 4753
 QY 3721 ATTTGAAGTGAATTAAGCAATCATGAGTCTGCTTAACTGAGAGCGGCTGCTGTAATTAAT 3780
 Db 4754 ATTTGAAGTGAATTAAGCAATCATGAGTCTGCTTAACTGAGAGCGGCTGCTGTAATTAAT 4813
 QY 3781 ATTTGAAGTGAATTAAGCAATCATGAGTCTGCTTAACTGAGAGCGGCTGCTGTAATTAAT 3840
 Db 4814 ATTTGAAGTGAATTAAGCAATCATGAGTCTGCTTAACTGAGAGCGGCTGCTGTAATTAAT 4873
 QY 3841 AAAAGTAAATGCTTTGAGTGAAGAACTCAAGTAGCAGAGAGCAGGTTTGGAAAGCT 3900
 Db 4874 AAAAGTAAATGCTTTGAGTGAAGAACTCAAGTAGCAGAGAGCAGGTTTGGAAAGCT 4933
 QY 3901 AAAAGTGAAGAGCAAGAAAAATTTCTATTGGGAATCTTCTGCTTCTATCACTTA 3960
 Db 4934 AAAAGTGAAGAGCAAGAAAAATTTCTATTGGGAATCTTCTGCTTCTATCACTTA 4993
 QY 3961 GTGAACAATGAAGTTTCTGCAAAATCAGAAAAATTAATCAGTAGCAGAGAAATCTGAAGC 4020
 Db 4994 GTGAACAATGAAGTTTCTGCAAAATCAGAAAAATTAATCAGTAGCAGAGAAATCTGAAGC 5053
 QY 4021 CAAAAATGGAATGTTGATGTCATGCTTATCAAGCGGACACCCAAAGTACAGAGACTTTA 4080
 Db 5054 CAAAAATGGAATGTTGATGTCATGCTTATCAAGCGGACACCCAAAGTACAGAGACTTTA 5113
 QY 4081 AATTTCAAGCTGGAAGTCAATATGAATCTGAGGGCTACTGTAATGAGTGAAGT 4140
 Db 5114 AATTTCAAGCTGGAAGTCAATATGAATCTGAGGGCTACTGTAATGAGTGAAGT 5173
 QY 4141 AACAACAAGTAAATGCTTCTATTAGTGGTGGAGATTAATCAATGAGTAAATGAGCGGAC 4200
 Db 5174 AACAACAAGTAAATGCTTCTATTAGTGGTGGAGATTAATCAATGAGTAAATGAGCGGAC 5233
 QY 4201 GCAAAAGCTCTTTTGAACAACCTCAAGTGAATGCTGCAAGTGAACGAGGAGCAAT 4260
 Db 5234 GCAAAAGCTCTTTTGAACAACCTCAAGTGAATGCTGCAAGTGAACGAGGAGCAAT 5293
 QY 4261 AGTTCTGAAGCGGATTAAGAAATTAATCAAGGGCTGTTCTGTCAATTAAGTGAAT 4320
 Db 5294 AGTTCTGAAGCGGATTAAGAAATTAATCAAGGGCTGTTCTGTCAATTAAGTGAAT 5353
 QY 4321 GACGTGAAGCTAGGCTTGAATTAATCTTCAATCGAAGAGCTAATGAATCAATGTCATT 4380
 Db 5354 GACGTGAAGCTAGGCTTGAATTAATCTTCAATCGAAGAGCTAATGAATCAATGTCATT 5413
 QY 4381 GCCAAGATGTCAAGAAATCTGATCTAGCAAAAGAAATCAAGCTTTAATTAATGA 4440
 Db 5414 GCCAAGATGTCAAGAAATCTGATCTAGCAAAAGAAATCAAGCTTTAATTAATGA 5473
 QY 4441 AAAAGTAAATAATTTAGAAAGATCTGTGATTAATCAAGTGAATGCTTATTAAGC 4500
 Db 5474 AAAAGTAAATAATTTAGAAAGATCTGTGATTAATCAAGTGAATGCTTATTAAGC 5533
 QY 4501 AAGGAACAATTAAGAAAAAGCAAGAAAAAGAAAGAGAGCGGTATTTGTAATCTGCTTA 4560
 Db 5534 AAGGAACAATTAAGAAAAAGCAAGAAAAAGAAAGAGAGCGGTATTTGTAATCTGCTTA 5593
 QY 4561 TCGGTTGCTGAAGCGGATTAATCCGCTGAGAGATTAATGCTTCAATCTGTTAA 4620
 Db 5594 TCGGTTGCTGAAGCGGATTAATCCGCTGAGAGATTAATGCTTCAATCTGTTAA 5653
 QY 4621 AATTAATTTAAAGCAATTAAGTGAAGCAATTAAGAAAGCGGAGAGATTAATTAAT 4680
 Db 5654 AATTAATTTAAAGCAATTAAGTGAAGCAATTAAGAAAGCGGAGAGATTAATTAAT 5713
 QY 4681 GCGAAACATGTAAATGTGAAGCAAAATCACTACGTGTTGTGAATGCGGCTTCTGA 4740

D	5714	GCAGAACTGTAAATGTGAGGCAAAATCATCTAGTGTGTGAAATGCGGCTTCTGGA	5773	D	6794	GCTGAGTTGGAGCTGAGAGAGCTGAGTGACAAGTTCTGTAGCAGTGAAATGAGTTGTA	6853
O	4741	CTTGCTATCAGCAAAAGATGCTTTTCAGGAATGGATCTGGAGCATGGCAAGACTTATCA	4800	O	5821	AATATATCGATAGGAGAAATTAATCATGCAAAAAATCACTGGAGAGGAAAAATGCGAGTT	5880
D	5774	CTTGCTATCAGCAAAAGATGCTTTTCAGGAATGGATCTGGAGCATGGCAAGACTTATCA	5833	D	6854	AATATATCGATAGGAGAAATTAATCATGCAAAAAATCACTGGAGAGGAAAAATGCGAGTT	6913
O	4801	AATGACAGATTGCAAAAGTGATATAAGAGAAATTTCTGCTGATTCCTTAATGTGAAC	4860	O	5881	ATTACAGAGCTGATGCGGTAAATGCTAATTTATGCAAGAAACAGTGTCTGAGTGCGCCGT	5940
D	5834	AATGACAGATTGCAAAAGTGATATAAGAGAAATTTCTGCTGATTCCTTAATGTGAAC	5893	D	6914	ATTACAGAGCTGATGCGGTAAATGCTAATTTATGCAAGAAACAGTGTCTGAGTGCGCCGT	6973
O	4861	GCAATTAATTCATTTCTGGGGTGAATGTTGCGGAACCATTTGCCGGTTCCTTTCTACG	4920	O	5941	GCAGCATATGAGAGCTTCAACCAAGTGTGAATGAATTAACAGGATCTTACAAAGCATATGTA	6000
D	5894	GCAATTAATTCATTTCTGGGGTGAATGTTGCGGAACCATTTGCCGGTTCCTTTCTACG	5953	D	6974	GCAGCATATGAGAGCTTCAACCAAGTGTGAATGAATTAACAGGATCTTACAAAGCATATGTA	7033
O	4921	GCGGATGAGAGCTGTTTGGCAATTAATCTCTTAATTAATAAAGCTCTGCTTTGATTAACA	4980	O	6001	AAAGATCTACAGTGAATTGCTAAAGAAAGACAGATGATTAATTAATTAATTAATTAATTAAT	6060
D	5954	GCGGATGAGAGCTGTTTGGCAATTAATCTCTTAATTAATAAAGCTCTGCTTTGATTAACA	6013	D	7034	AAAGATCTACAGTGAATTGCTAAAGAAAGACAGATGATTAATTAATTAATTAATTAATTAAT	7093
O	4981	GGAACGAAAGTAAATCTTTAGTGAAGAAATACAAAAGTCAATGTACAAGCTTTGAAT	5040	O	6061	GTAATTAAGTGTAGATTAAGTATTCAAAATCTTAATATTAATTAATTAATTAATTAATTAAT	6120
D	6014	GGAACGAAAGTAAATCTTTAGTGAAGAAATACAAAAGTCAATGTACAAGCTTTGAAT	6073	D	7094	GTAATTAAGTGTAGATTAAGTATTCAAAATCTTAATATTAATTAATTAATTAATTAATTAAT	7153
O	5041	GATTCTCATATTACAAAGCTTTCTGCTGAGAGCGCTGCAAGTATTAAAGCGCTGGAATC	5100	O	6121	AAAAAGAAATTAAGTAAATTAAGAAAGATTTGTTACCAATAGTTCACTCACTTATTA	6180
D	6074	GATTCTCATATTACAAAGCTTTCTGCTGAGAGCGCTGCAAGTATTAAAGCGCTGGAATC	6133	D	7154	AAAAAGAAATTAAGTAAATTAAGAAAGATTTGTTACCAATAGTTCACTCACTTATTA	7213
O	5101	GGAGGAATGATCTGCTGCAATGATGTTCTGATGAAACGGAAGCTTTAGTTAGTAAATCT	5160	O	6181	AAATCTTTATTTGGCAATATGCGGCTGTTCAAGAACAGCGGAGTGGCAGGAAGCTTTAAT	6240
D	6134	GGAGGAATGATCTGCTGCAATGATGTTCTGATGAAACGGAAGCTTTAGTTAGTAAATCT	6193	D	7214	AAATCTTTATTTGGCAATATGCGGCTGTTCAAGAACAGCGGAGTGGCAGGAAGCTTTAAT	7273
O	5161	GAGTTTGAAGAGTAAAGTCTTTCAATGATGATGCAAAAAGATCAAAAAACATAATATCA	5220	O	6241	ATCAACAAGGTTTATGAGAAACAGAAAGCTTTGTAGAAAATTCATATTAATTAATTAATTAAT	6300
D	6194	GAGTTTGAAGAGTAAAGTCTTTCAATGATGATGCAAAAAGATCAAAAAACATAATATCA	6253	D	7274	ATCAACAAGGTTTATGAGAAACAGAAAGCTTTGTAGAAAATTCATATTAATTAATTAATTAAT	7333
O	5221	ATTGCCGGAATGCAATGAGAGAAAGCGGCTGGAAGTTTGGAGCAACAGTTGCTCATACA	5280	O	6301	CATTATTTCTGTAATTAATCAGAGATTAACAGAAATTCATCGAGATGATAGTTCTGTTGGT	6360
D	6254	ATTGCCGGAATGCAATGAGAGAAAGCGGCTGGAAGTTTGGAGCAACAGTTGCTCATACA	6313	D	7334	CATTATTTCTGTAATTAATCAGAGATTAACAGAAATTCATCGAGATGATAGTTCTGTTGGT	7393
O	5281	AATATTGGAAAAACATCAGTTTACGTTATTTGTAATAAACAGTAAATTAACAACGCGGAT	5340	O	6361	GTTGGTGGAAATGTAAGAGTGAAGAGCTTCTTCTGATACCAATTAATTAATTAATTAATTAAT	6420
D	6314	AATATTGGAAAAACATCAGTTTACGTTATTTGTAATAAACAGTAAATTAACAACGCGGAT	6373	D	7394	GTTGGTGGAAATGTAAGAGTGAAGAGCTTCTTCTGATACCAATTAATTAATTAATTAATTAAT	7453
O	5341	GATCAGATAGAAAAATATCAATGATGATGCAAAAAGATTAATATGATGCAATATCTATA	5400	O	6421	AAAGCAAGAGTTTGAAGAAACATCAATGATCTGATAGAGTTTCCGAGAGAAAGCTGAAAT	6480
D	6374	GATCAGATAGAAAAATATCAATGATGATGCAAAAAGATTAATATGATGCAATATCTATA	6433	D	7454	AAAGCAAGAGTTTGAAGAAACATCAATGATCTGATAGAGTTTCCGAGAGAAAGCTGAAAT	7513
O	5401	GCAAGTCGAGTTGAGAGAGCAAAAGAGCTTCTGCAAGAGAGCTTCTGCAAGTACTAAC	5460	O	6481	ACAGCAGATTCTAAGCAAGAAATTTCTTTTGGAGTCGAGTCGACAGCGCGGTA	6540
D	6434	GCAAGTCGAGTTGAGAGAGCAAAAGAGCTTCTGCAAGAGAGCTTCTGCAAGTACTAAC	6493	D	7514	ACAGCAGATTCTAAGCAAGAAATTTCTTTTGGAGTCGAGTCGACAGCGCGGTA	7573
O	5461	TTGAATTAAGACAGTTTCTTCTCAATGTAACAACATGATTAATGAACAAAGATTTAAGGA	5520	O	6541	GGAGCCGAGAGTGGAGAGAACGCTTCCGTTAAATGAATTTGCAAGAGAAAGACGGAAGTGAAT	6600
D	6494	TTGAATTAAGACAGTTTCTTCTCAATGTAACAACATGATTAATGAACAAAGATTTAAGGA	6553	D	7574	GGAGCCGAGAGTGGAGAGAACGCTTCCGTTAAATGAATTTGCAAGAGAAAGACGGAAGTGAAT	7633
O	5521	GAAAAATTAATGAGAAATGAGAAAGGCAAAATGTTAATGTTCTAGCTGAATAATCGAGTCA	5580	O	6601	GTTGAGAGAGCAAGAAATTTTGTGTAATAAAGCTGAGATTAACAGCAAAAAGCTTATAGTTCT	6660
D	6554	GAAAAATTAATGAGAAATGAGAAAGGCAAAATGTTAATGTTCTAGCTGAATAATCGAGTCA	6613	D	7634	GTTGAGAGAGCAAGAAATTTTGTGTAATAAAGCTGAGATTAACAGCAAAAAGCTTATAGTTCT	7693
O	5581	GTTGTCACAAATGCAACAGTGTCTTCCGAGCAAGTGAACAAAGCTGCAAGTGAAGAGCTGGA	5640	O	6661	GTTGCAATTTGGAATATGCCAGTGCAGAGTGGCTGCAAAAAGAGCTGCAATTTGAGACAGCA	6720
D	6614	GTTGTCACAAATGCAACAGTGTCTTCCGAGCAAGTGAACAAAGCTGCAAGTGAAGAGCTGGA	6673	D	7694	GTTGCAATTTGGAATATGCCAGTGCAGAGTGGCTGCAAAAAGAGCTGCAATTTGAGACAGCA	7753
O	5641	GTAAGAGTTAATTAATAATTAACAATAATCTTCTGCAACATTAATAAATAGTACTAAAT	5700	O	6721	GTTGCAATTTGGAATATGCCAGTGCAGAGTGGCTGCAAAAAGAGCTGCAATTTGAGACAGCA	6780
D	6674	GTAAGAGTTAATTAATAATTAACAATAATCTTCTGCAACATTAATAAATAGTACTAAAT	6733	D	7754	GTTGCAATTTGGAATATGCCAGTGCAGAGTGGCTGCAAAAAGAGCTGCAATTTGAGACAGCA	7813
O	5701	GTAAGAGTTAATTAATAATTAACAATAATCTTCTGCAACATTAATAAATAGTACTAAAT	5760	D	6781	ACTCGAAACAGATTAGATGATATAGCAGAAATGAGATTAATTAATTAATTAATTAATTAAT	6840
D	6734	GTAAGAGTTAATTAATAATTAACAATAATCTTCTGCAACATTAATAAATAGTACTAAAT	6793	D	7814	ACTCGAAACAGATTAGATGATATAGCAGAAATGAGATTAATTAATTAATTAATTAATTAAT	7873
O	5761	GCTGAGTTGAGCTGAGAGAGCTGAGTGAACAGTTCTGTAGCAGTGAATTAAGATTGTA	5820	O	6841	TCAGCCGAGAGCTGGAATTTCTTGACAGCCGAGATTAATCTGAGTGGTTTCTGTCAATTAATTT	6900
D				D	7874	TCAGCCGAGAGCTGGAATTTCTTGACAGCCGAGATTAATCTGAGTGGTTTCTGTCAATTAATTT	7933

QY 6901 GCAAAATTAAGTGAAGACAGATATGCATAGTACTTTACACTCTTTCTATCTGATGTAAT 6960
|
|
|
Db 7934 GCAAAATTAAGTGAAGACAGATATGCATAGTACTTTACACTCTTTCTATCTGATGTAAT 7993
|
|
|
QY 6961 GTTAAAGCTCTTAATTAATTAATTTGAAATTCCTTTGACAGCGGCGGAGAGCGCAGGCTCT 7020
|
|
|
|
Db 7994 GTTAAAGCTCTTAATTAATTAATTTGAAATTCCTTTGACAGCGGCGGAGAGCGCAGGCTCT 8053
|
|
|
|
QY 7021 GCAGCAGTTTACCGGAGTGGTTCTGTTTACATATAAATAGTTCTGTATAGCTGAGTT 7080
|
|
|
|
Db 8054 GCAGCAGTTTACCGGAGTGGTTCTGTTTACATATAAATAGTTCTGTATAGCTGAGTT 8113
|
|
|
|
QY 7081 CACATTAATCTGATTTGACTTCCTGTACGAGAAAAAGTAAATGTAAACGCGCAAAAGAGAA 7140
|
|
|
|
Db 8114 CACATTAATCTGATTTGACTTCCTGTACGAGAAAAAGTAAATGTAAACGCGCAAAAGAGAA 8173
|
|
|
|
QY 7141 AAAAAATTTAAGCAACAGCAGCAAAATGAGGAATCGGAGGAGCAATCGGAGCCAT 7200
|
|
|
|
Db 8174 AAAAAATTTAAGCAACAGCAGCAAAATGAGGAATCGGAGGAGCAATCGGAGCCAT 8233
|
|
|
|
QY 7201 GTCTTGTAATTAATTTTGGAAACAGCTGTAGAAAGATAGAAAAATTTGAAAGAAAAAGA 7260
|
|
|
|
Db 8234 GTCTTGTAATTAATTTTGGAAACAGCTGTAGAAAGATAGAAAAATTTGAAAGAAAAAGA 8293
|
|
|
|
QY 7261 ACAGAAAGTTTAAAACTTTAGACAGAGTTAACAAAGAACAGATTAATTAATGAT 7320
|
|
|
|
Db 8294 ACAGAAAGTTTAAAACTTTAGACAGAGTTAACAAAGAACAGATTAATTAATGAT 8353
|
|
|
|
QY 7321 GCTAGAAAAAAATCTTCAATCAGAGGATTTTCTACAGAGATCTTCTGTAAGAGCG 7380
|
|
|
|
Db 8354 GCTAGAAAAAAATCTTCAATCAGAGGATTTTCTACAGAGATCTTCTGTAAGAGCG 8413
|
|
|
|
QY 7381 GATAGAGGAGATATCTCAGGAGAGAGAAATTAAGCCATTGTGAGAGCTTCTGATTAAT 7440
|
|
|
|
Db 8414 GATTAAGAGGATATCTCAGGAGAGAGAAATTAAGCCATTGTGAGAGCTTCTGATTAAT 8473
|
|
|
|
QY 7441 GAAAAAAATGTAGATTAATTAACACAGACACAGAAATATATATCATCTTCTGAGGTTTG 7500
|
|
|
|
Db 8474 GAAAAAAATGTAGATTAATTAACACAGACACAGAAATATATATCATCTTCTGAGGTTTG 8533
|
|
|
|
QY 7501 GGAATCGCAGGCTTGTCCGATCAGAGAACAGTGCAGTTTACAATTTTAAAAAGAAAT 7560
|
|
|
|
Db 8534 GGAATCGCAGGCTTGTCCGATCAGAGAACAGTGCAGTTTACAATTTTAAAAAGAAAT 8593
|
|
|
|
QY 7561 TCCGAGGTTACTGTGAAAAATCTTTTGTGAAAGCAGCTGAAAAAGTAAATGTAGATCG 7620
|
|
|
|
Db 8594 TCCGAGGTTACTGTGAAAAATCTTTTGTGAAAGCAGCTGAAAAAGTAAATGTAGATCG 8653
|
|
|
|
QY 7621 GATATTACAGAAATGTTGCTTTAAACAGCATATCAAGGTCTGTAGAGCAATTTGGAAAT 7680
|
|
|
|
Db 8654 GATATTACAGAAATGTTGCTTTAAACAGCATATCAAGGTCTGTAGAGCAATTTGGAAAT 8713
|
|
|
|
QY 7681 GGAAGTCCCTATGACAGAAATTAATTTCTATGGAAGATCAAAATATCATCTTAAAAATTTCT 7740
|
|
|
|
Db 8714 GGAAGTCCCTATGACAGAAATTAATTTCTATGGAAGATCAAAATATCATCTTAAAAATTTCT 8773
|
|
|
|
QY 7741 AAGCTATTAGAAAAATATATGATGTTATTTGTPAAAAGATTAATCGGAATTTGAGACGGAA 7800
|
|
|
|
Db 8774 AAGCTATTAGAAAAATATATGATGTTATTTGTPAAAAGATTAATCGGAATTTGAGACGGAA 8833
|
|
|
|
QY 7801 GCAAAAAGATTAACCGTAGAGAGCGGTAGCTGCCGAGCCATTAATCTCAAAAAGCAAGAAAT 7860
|
|
|
|
Db 8834 GCAAAAAGATTAACCGTAGAGAGCGGTAGCTGCCGAGCCATTAATCTCAAAAAGCAAGAAAT 8893
|
|
|
|
QY 7861 GAAAATGAATTCAGAGGTTGAAAATTTGAGAGAGATTTTCAATGAAGAAAAATGAGTAATCT 7920
|
|
|
|
Db 8894 GAAAATGAATTCAGAGGTTGAAAATTTGAGAGAGATTTTCAATGAAGAAAAATGAGTAATCT 8953
|
|
|
|
QY 7921 AGCCCTTTCAAGCAATTTGGAAGAGAAATCAATGTCAAAAGTGAAAAAGAAAAACAAGATG 7980
|
|
|
|
Db 8954 AGCCCTTTCAAGCAATTTGGAAGAGAAATCAATGTCAAAAGTGAAAAAGAAAAACAAGATG 9013
|
|
|
|

QY 7981 ACTGCTGAATCTCAAGAGAGCTTCTGTAGAGACAGTAGAGGCGGACAGAAATTAATTTCCGAA 8040
|
|
|
Db 9014 ACTGCTGAATCTCAAGAGAGCTTCTGTAGAGACAGTAGAGGCGGACAGAAATTAATTTCCGAA 9073
|
|
|
|
QY 8041 GCAAAAAGATGCGGAAGCTCTTAATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100
|
|
|
|
Db 9074 GCAAAAAGATGCGGAAGCTCTTAATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 9133
|
|
|
|
QY 8101 CATGCAATTAATGTGAAATATGGAAGCAACATATAAAATGAAGTAAACGCGATTTCTAAA 8160
|
|
|
|
Db 9134 CATGCAATTAATGTGAAATATGGAAGCAACATATAAAATGAAGTAAACGCGATTTCTAAA 9193
|
|
|
|
QY 8161 GCAGTAAACAGGTTCTGTATTGGAGAGAGTTGAGTCAACCAAGCGAAGCTTACTGCTCA 8220
|
|
|
|
Db 9194 GCAGTAAACAGGTTCTGTATTGGAGAGAGTTGAGTCAACCAAGCGAAGCTTACTGCTCA 9253
|
|
|
|
QY 8221 GGTAAACCTATGATGAGAAAGTTGAGGAGAAATTTGTTCAAGAACAAATCGATTGAATGCA 8280
|
|
|
|
Db 9254 GGTAAACCTATGATGAGAAAGTTGAGGAGAAATTTGTTCAAGAACAAATCGATTGAATGCA 9313
|
|
|
|
QY 8281 AATTTCTAAAGTGAAGGTTTGGATGCAATTAAGTAACTGCTTAAATCTTCTGATGATCA 8340
|
|
|
|
Db 9314 AATTTCTAAAGTGAAGGTTTGGATGCAATTAAGTAACTGCTTAAATCTTCTGATGATCA 9373
|
|
|
|
QY 8341 GGAATGAGAGAGGAATTTGCCGAGCAGAGTGAATATCTTCTACAGACAAAGTAAATCT 8400
|
|
|
|
Db 9374 GGAATGAGAGAGGAATTTGCCGAGCAGAGTGAATATCTTCTACAGACAAAGTAAATCT 9433
|
|
|
|
QY 8401 GAATCCGTAGTTCGTTTACGAAAGCAGATTAATGAAATTAATGATTACACAAAAAAATAT 8460
|
|
|
|
Db 9434 GAATCCGTAGTTCGTTTACGAAAGCAGATTAATGAAATTAATGATTACACAAAAAAATAT 9493
|
|
|
|
QY 8461 ATTTCAAGAGTCAATGCTCTTCTGTTAAATGATACAAAGAAATGAAGCCGAAATTAATCT 8520
|
|
|
|
Db 9494 ATTTCAAGAGTCAATGCTCTTCTGTTAAATGATACAAAGAAATGAAGCCGAAATTAATCT 9553
|
|
|
|
QY 8521 TTAGCGGTAGCCGCGTGCATGACCAAGGAACAAACAAAGCATTTACGAGATCAAAACAAG 8580
|
|
|
|
Db 9554 TTAGCGGTAGCCGCGTGCATGACCAAGGAACAAACAAAGCATTTACGAGATCAAAACAAG 9613
|
|
|
|
QY 8581 TTAACTTCTACAACTGTAAATGAGAGAAACGTATCTCAACTTGTGCAAAAACCTTTGGCT 8640
|
|
|
|
Db 9614 TTAACTTCTACAACTGTAAATGAGAGAAACGTATCTCAACTTGTGCAAAAACCTTTGGCT 9673
|
|
|
|
QY 8641 AAAAAATGAATTTATGAAATGTAAAGAACTGAGAGAGCTTACTGAGACGGAGAAACA 8700
|
|
|
|
Db 9674 AAAAAATGAATTTATGAAATGTAAAGAACTGAGAGAGCTTACTGAGACGGAGAAACA 9733
|
|
|
|
QY 8701 GAGGCCGTTGAAAAATTAACAAAGATTAACAGAGCAATTTGTTGACAGAAATTTGGGAA 8760
|
|
|
|
Db 9734 GAGGCCGTTGAAAAATTAACAAAGATTAACAGAGCAATTTGTTGACAGAAATTTGGGAA 9793
|
|
|
|
QY 8761 ATTTGAGATTAATTTAGAAAACGATTCAGAGATTAATCGATTTGTAAGTCAACGAGAC 8820
|
|
|
|
Db 9794 ATTTGAGATTAATTTAGAAAACGATTCAGAGATTAATCGATTTGTAAGTCAACGAGAC 9853
|
|
|
|
QY 8821 GGAACCAAGAGAGTCTTGTCCGAAAAGATGATTTTCTGTGAAAAATCAATTTCAAGG 8880
|
|
|
|
Db 9854 GGAACCAAGAGAGTCTTGTCCGAAAAGATGATTTTCTGTGAAAAATCAATTTCAAGG 9913
|
|
|
|
QY 8881 GAAACAAAATCATCCATTTGAAGATTAAGCCAGAAATTTGTAAGCCGGAAGTGAATGTA 8940
|
|
|
|
Db 9914 GAAACAAAATCATCCATTTGAAGATTAAGCCAGAAATTTGTAAGCCGGAAGTGAATGTA 8973
|
|
|
|
QY 8941 GATGCTTTGAATGAACCTGATGATGATCTTACAGGAAAAAGTGTGCTATGCTGGAATTT 9000
|
|
|
|
Db 9974 GATGCTTTGAATGAACCTGATGATGATCTTACAGGAAAAAGTGTGCTATGCTGGAATTT 10033
|
|
|
|
QY 9001 GGTATTGGAATGTTGATGATTAATATGATGATTAAGAAAAATGTAGAACCCAAAAATCGGA 9060
|
|
|
|
Db 10034 GGTATTGGAATGTTGATGATTAATATGATGATTAAGAAAAATGTAGAACCCAAAAATCGGA 10093
|
|
|
|
QY 9061 AGACATGCTAATTTGTGAAACCTACGGAAGAAACAAGAAATATCAAGCAATTTCAAGAGCAAAA 9120
|
|
|
|

Db	1021	AATCATATAATATCTGCTTCTGTGGCAATAGAGAGAGTGGAAAGACTTTTCCAGGAGTG	1080
Oy	1999	GAAGGAGTAATGTAAAGGCATTAAATGAAGCTCAAAATCTTGAGGCACTACGTCAAGT	2058
Db	1081	GAAGGAGTAATGTAAAGGCATTAAATGAAGCTCAAAATCTTGAGGCACTACGTCAAGT	1140
Oy	2059	GGAAAGTCGTGCTGTAACGAAGGAAGAAAAAGAAACTTATTTGGAAATGACAGAGTTTTT	2118
Db	1141	GGAAAGTCGTGCTGTAACGAAGGAAGAAAAAGAAACTTATTTGGAAATGACAGAGTTTTT	1200
Oy	2119	TATGGAAACTATAAAAAATTAATGCTTCTGTGCAATTCGCCATCATGCTGAATTTGATCG	2178
Db	1201	TATGGAAACTATAAAAAATTAATGCTTCTGTGCAATTCGCCATCATGCTGAATTTGATCG	1260
Oy	2179	GAAGGAAAAATTTGATATCAACAGTGAATAATTAATTAATAAATTCCTTCAAAAATG	2238
Db	1261	GAAGGAAAAATTTGATATCAACAGTGAATAATTAATTAATAAATTCCTTCAAAAATG	1320
Oy	2239	GCAAAAGTCGTATTTGATTAATTTGAACTTTTAAAGAGACTTTTGGAAAAAGAACAA	2298
Db	1321	GCAAAAGTCGTATTTGATTAATTTGAACTTTTAAAGAGACTTTTGGAAAAAGAACAA	1380
Oy	2299	ACTCCAGAAATATGATCCGAAGAATTAATGAATCTATTGAAAAATTAATGAATGACTTTCA	2358
Db	1381	ACTCCAGAAATATGATCCGAAGAATTAATGAATCTATTGAAAAATTAATGAATGACTTTCA	1440
Oy	2359	GAAAAAATTTGATGAAAAACCGAGCTTTTACTTAATGGTGAAGAAATGACATTAATCTT	2418
Db	1441	GAAAAAATTTGATGAAAAACCGAGCTTTTACTTAATGGTGAAGAAATGACATTAATCTT	1500
Oy	2419	CCGAGTGGAACTTCAAAAAACAGAACTGCTATAGAAATTTGCAACTATGTTCAAGGAGAA	2478
Db	1501	CCGAGTGGAACTTCAAAAAACAGAACTGCTATAGAAATTTGCAACTATGTTCAAGGAGAA	1560
Oy	2479	ATGAAAAAATTAAGAGGAAAAATTCACGAAGAGATTTTAAAGCTTTTCAAGAGATAGT	2538
Db	1561	ATGAAAAAATTAAGAGGAAAAATTCACGAAGAGATTTTAAAGCTTTTCAAGAGATAGT	1620
Oy	2539	GGACTGATTAAGAAGAACTTTGAATTTTACAGAGTAGAAGAAATTAAGCAAAATTTTCAACT	2598
Db	1621	GGACTGATTAAGAAGAACTTTGAATTTTACAGAGTAGAAGAAATTAAGCAAAATTTTCAACT	1680
Oy	2599	TTTACCTCTTCCGAGCTTAATGAGAAAGAGATGTTCTTTGTGGGAGAGCTGTTTCG	2658
Db	1681	TTTACCTCTTCCGAGCTTAATGAGAAAGAGATGTTCTTTGTGGGAGAGCTGTTTCG	1740
Oy	2659	TGGGTAGAACAGAGAAATTAAGCAAGGTATCCGTGGAAAAAGAGCTAAACCTTGCTGCA	2718
Db	1741	TGGGTAGAACAGAGAAATTAAGCAAGGTATCCGTGGAAAAAGAGCTAAACCTTGCTGCA	1800
Oy	2719	AAAAAAGATTTAAATATATTAAGACTATCAATTAAGACAGAAACAGTGAATTTTATGTTGAAT	2778
Db	1801	AAAAAAGATTTAAATATATTAAGACTATCAATTAAGACAGAAACAGTGAATTTTATGTTGAAT	1860
Oy	2779	ATTGGACTTGGAGAGAGCAATACATCCGGAAGTGCAGTCCGAGAGAAATTAATGTTCAA	2838
Db	1861	ATTGGACTTGGAGAGAGCAATACATCCGGAAGTGCAGTCCGAGAGAAATTAATGTTCAA	1920
Oy	2839	AGATGGAATAATTCAGCTATTCGTAGAACCTAAAGAAAAAGCTGAATTAATCAGAGAAAAAT	2898
Db	1921	AGATGGAATAATTCAGCTATTCGTAGAACCTAAAGAAAAAGCTGAATTAATCAGAGAAAAAT	1980
Oy	2899	ATTAAATGAGATGCAATTAAGCAAGCTTTTTCATGTAAGCGGAGCTTTTAAATGATGAGCTCA	2958
Db	1981	ATTAAATGAGATGCAATTAAGCAAGCTTTTTCATGTAAGCGGAGCTTTTAAATGATGAGCTCA	2040
Oy	2959	GGTGGAAATGCAATTAAGCAATGGAATGGGAAGTTATAGTGAAGGTATCAGTAAGCAAGCTT	3018
Db	2041	GGTGGAAATGCAATTAAGCAATGGAATGGGAAGTTATAGTGAAGGTATCAGTAAGCAAGCTT	2100
Oy	3019	TCCATTGATGACGAAGCAATTTTGAAGCTTAATAAAAAATTCCTTTAAACAGTAAGAT	3078
Db	2101	TCCATTGATGACGAAGCAATTTTGAAGCTTAATAAAAAATTCCTTTAAACAGTAAGAT	2160

QY	3079	GATACCTTCTGTTTGGAAAGCTCCCGGTTCAAGCGGAAATCGGAACGAAAAATGCGGCGGTC	3138
Db	2161	GATACCTTCTGTTTGGAAAGCTCCCGGTTCAAGCGGAAATCGGAACGAAAAATGCGGCGGTC	2220
QY	3139	GGGGTGTGCTGTGCGGTAATGATTTATTTCAACCAAAAGCTTCCATTGAAGATNAT	3198
Db	2221	GGGGTGTGCTGTGCGGTAATGATTTATTTCAACCAAAAGCTTCCATTGAAGATNAT	2280
QY	3199	GACGAGGACAAAGTAAATGATTAAGATTAAGATGATGAAGTAAACATTAATCGCGAA	3258
Db	2281	GACGAGGACAAAGTAAATGATTAAGATTAAGATGATGAAGTAAACATTAATCGCGAA	2340
QY	3259	TCTTTAGAAGTATGATGCAAAAAACGACCGGAACATCAAGTATTTCTGTGCGGAGGA	3318
Db	2341	TCTTTAGAAGTATGATGCAAAAAACGACCGGAACATCAAGTATTTCTGTGCGGAGGA	2400
QY	3319	ATTAATAAGTGTGAAGTAAACCGAGTGAAGAAAAACGAAATCAGAAAGAAAGACAGAG	3378
Db	2401	ATTAATAAGTGTGAAGTAAACCGAGTGAAGAAAAACGAAATCAGAAAGAAAGACAGAG	2460
QY	3379	GGATTTTGTGCGAAATTCGGAACAAAGTGAATCTGTATAAAAAATTAACGATAGT	3438
Db	2461	GGATTTTGTGCGAAATTCGGAACAAAGTGAATCTGTATAAAAAATTAACGATAGT	2520
QY	3439	ATGATTAATTAACGAAAAAATTAACAATTACATTTCTGAAGAGTAAAAAAGCGGG	3498
Db	2521	ATGATTAATTAACGAAAAAATTAACAATTACATTTCTGAAGAGTAAAAAAGCGGG	2580
QY	3499	AATCTTCTTTCGACGTTTCTCATCTCCGATTAAGAGACCGTCTTTCAGTTGGAGCT	3558
Db	2581	AATCTTCTTTCGACGTTTCTCATCTCCGATTAAGAGACCGTCTTTCAGTTGGAGCT	2640
QY	3559	TCTGGAAGTGTCTTCTTCAATATATTAAGAAACATCTGCTGCTGATGAGTGA	3618
Db	2641	TCTGGAAGTGTCTTCTTCAATATATTAAGAAACATCTGCTGCTGATGAGTGA	2700
QY	3619	AAGATAAATTTGAAGGAGCAATAAAGAGTAGAGTGAATTTCTGATTTCTTCTT	3678
Db	2701	AAGATAAATTTGAAGGAGCAATAAAGAGTAGAGTGAATTTCTGATTTCTTCTT	2760
QY	3679	GTTGAGCATGGGGCGGATC 3698	
Db	2761	GTTGAGCATGGGGCGGATC 2780	
RESULT 6			
US-10-647-057-10			
; Sequence 10, Application US/10647057			
; Publication No. US20040047871A1			
GENERAL INFORMATION:			
APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION			
APPLICANT: NAKARADA, T			
APPLICANT: STEWART, GEORGE			
APPLICANT: NARAYANAN, SANJEEV			
APPLICANT: CHENAPPA, M.			
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP			
TITLE OF INVENTION: THREEOF			
FILE REFERENCE: PCT-30962			
CURRENT APPLICATION NUMBER: US/10/647, 057			
CURRENT FILING DATE: 2003-08-22			
NUMBER OF SEQ ID NOS: 15			
SOFTWARE: Patent version 3.0			
SEQ ID NO 10			
LENGTH: 2780			
TYPE: DNA			
ORGANISM: Fusobacterium necrophorum			
US-10-647-057-10			
Query Match			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

919 GGGAGATTTATTAAGAAAAGAGGAGAAAAGAACTTATTAACACTCTTTAAAGTTATCA 978
1 GGGAAATTTATTAAGAAAAGAGGAGAAAAGAACTTATTAACACTCTTTAAAGTTATCA 60
979 GATGTGAAGCTTCCGTAAAGCTAATAAGAAAAGTCATGGAAGAAAGTTACATT 1038
61 GATGTGAAGCTTCCGTAAAGCTAATAAGAAAAGTCATGGAAGAAAGTTACATT 120
1039 ACAGCTGAAGCAAAAGATTTCTATGATGCAACTTATAGTTACTAGCTTGCAAAAGCACTCT 1098
121 ACAGCTGAAGCAAAAGATTTCTATGATGCAACTTATAGTTACTAGCTTGCAAAAGCACTCT 180
1099 TTTAGCTTTGTTACAGGTTCTAATTTCTCTATCAATTTAAATGAAATTTTATAGTTATG 1158
181 TTTAGCTTTGTTACAGGTTCTAATTTCTCTATCAATTTAAATGAAATTTTATAGTTATG 240
1159 ACAATTAAGTCCAGTGTGCTTATTTGAAAAGATGCCAAAGTGAAGCAACAGAGAAAG 1218
241 ACAGCTAAGTCCAGTGTGCTTATTTGAAAAGATGCCAAAGTGAAGCAACAGAGAAAG 300
1219 GCAAAATTTCTATTCTTACAGTGAAGTAAAGCAACTAGGAGCAGCTACTTCTCATTA 1278
301 GCAAAATTTCTATTCTTACAGTGAAGTAAAGCAACTAGGAGCAGCTACTTCTCATTA 360
1279 AAAATTTACCAATTTATATTTGAGAAAGCCAAATGAAAACTTCTCAGTATCGAGCGGGA 1338
361 AAAATTTACCAATTTATATTTGAGAAAGCCAAATGAAAACTTCTCAGTATCGAGCGGGA 420
1339 TATATTTCTGCAAAAAGTAACTTCCAATGTAACTATTTGAAGGAAGTAAATTCGAAGGA 1398
421 TATATTTCTGCAAAAAGTAACTTCCAATGTAACTATTTGAAGGAAGTAAATTCGAAGGA 480
1399 AGAGCAGATATTAATCTTCAAAATCTGAAAAATCTATTTAGTCTTCTGTTTCTGTTGAAG 1458
481 AGAGCAGATATTAATCTTCAAAATCTGAAAAATCTATTTAGTCTTCTGTTTCTGTTGAAG 540
1459 ATGAGAGATTCCAATTAAGTAGCTCTTTCAGTATTTGTTGGAAGCAAGAAAAATTAATCT 1518
541 ATGAGAGATTCCAATTAAGTAGCTCTTTCAGTATTTGTTGGAAGCAAGAAAAATTAATCT 600
1519 TCCGTCAGAGTTGCTTAAAGAGCAAAAGTAAATCAGAAACGAGATGTAATTTGGA 1578
601 TCCGTCAGAGTTGCTTAAAGAGCAAAAGTAAATCAGAAACGAGATGTAATTTGGA 660
1579 AGTGAAGCGATTAATCTTCAATTCGAGCTGCTGTAAGAGGTGATTTGGGGATATGTGTAAT 1638
661 AGTGAAGCGATTAATCTTCAATTCGAGCTGCTGTAAGAGGTGATTTGGGGATATGTGTAAT 720
1639 GGGGTTGTGCTGCAAAATATTTCTAATTAATGCTTCTCCCGTATAGATGTAGATGA 1698
721 GGGGTTGTGCTGCAAAATATTTCTAATTAATGCTTCTCCCGTATAGATGTAGATGA 780
1699 TATCTACATGCCAAGAGCACTAATATGTGAGGCTCATTAATTAATTAATAGTGT 1758
781 TATCTACATGCCAAGAGCACTAATATGTGAGGCTCATTAATTAATTAATAGTGT 840
1759 CTGCAAAAGAGATCTGATTTGGGAATTCGAAGTTATAGATGATCAGGTTATGATCA 1818
841 CTGCAAAAGAGATCTGATTTGGGAATTCGAAGTTATAGATGATCAGGTTATGATCA 900
1819 GGTCTATCTTAAATCAATTTTATAGATCAATTAAGCAAGCGGTTTGAAGAGACAGTCAAT 1878
901 GGTCTATCTTAAATCAATTTTATAGATCAATTAAGCAAGCGGTTTGAAGAGACAGTCAAT 960
1879 GAGGAAATTAAGAAATTAAGCTAAGAACTTATTAATGTCGTGTCTGCAACCTATAGA 1938
961 GAGGAAATTAAGAAATTAAGCTAAGAACTTATTAATGTCGTGTCTGCAACCTATAGA 1020
1939 AATCATTAATTAATTTGCTTCTGTGCAATTAAGAGAGAGTGAAGACTTTTCTCAGAGAG 1998
1021 AATCATTAATTAATTTGCTTCTGTGCAATTAAGAGAGAGTGAAGACTTTTCTCAGAGAG 1080
1999 GAAAGGAGTAAATGTAAAGGCAATTAATGAAGCTCAAAATCTTCGAGCGCATACGTCAGT 2058

1081 GAAAGGAGTAAATGTAAAGGCAATTAATGAAGCTCAAAATCTTCGAGCGCATACGTCAGT 1140
2059 GGAAGTGTGCTGTATCCGAAAGAGAAAGAAAAGAACTTATTTGAATATGCAAGCTTTT 2118
1141 GGAAGTGTGCTGTATCCGAAAGAGAAAGAAAAGAACTTATTTGAATATGCAAGCTTTT 1200
2119 TATGAAAATCTATTAATAATTAATGCTTCTGTGCAATTTGCCATTCATGCTGAATGGTATCG 2178
1201 TATGAAAATCTATTAATAATTAATGCTTCTGTGCAATTTGCCATTCATGCTGAATGGTATCG 1260
2179 GAAAGAAAATTTGATATCAACAGTGAATAATTAATAATTAATAATCTTCAAAAATG 2238
1261 GAAAGAAAATTTGATATCAACAGTGAATAATTAATAATTAATAATCTTCAAAAATG 1320
2239 GCAAGCTCTGTTATTTGATTAATTTAGAACTTTTAAAGAGCTTTTGGAAAAGAAACGAA 2298
1321 GCAAGCTCTGTTATTTGATTAATTTAGAACTTTTAAAGAGCTTTTGGAAAAGAAACGAA 1380
2299 ACTCCAGAAATATGATCCGAAAGATATTGAATCTATTTGAAAAATTAATGAATGCAATTTTCA 2358
1381 ACTCCAGAAATATGATCCGAAAGATATTGAATCTATTTGAAAAATTAATGAATGCAATTTTCA 1440
2359 GAAAAATTTGATGGAAGAAACCGAGCTTTTACTAAATGTTGAAGAAATGACAAATTAATCTT 2418
1441 GAAAAATTTGATGGAAGAAACCGAGCTTTTACTAAATGTTGAAGAAATGACAAATTAATCTT 1500
2419 CCGATGGAATCTTCAAAAACAGGAACTGCTATTAAGAAATTTGAAACTATGTTCAAGGAGAA 2478
1501 CCGATGGAATCTTCAAAAACAGGAACTGCTATTAAGAAATTTGAAACTATGTTCAAGGAGAA 1560
2479 ATGAAAAAATTTAGAGAGAAAAATTAACGAAAGATTTTAAAGCTTTTTCAGAAAGATTTGAGT 2538
1561 ATGAAAAAATTTAGAGAGAAAAATTAACGAAAGATTTTAAAGCTTTTTCAGAAAGATTTGAGT 1620
2539 GGAATGATTAAGAAACTTTGAAATTTTACAGAGTGAAGAAATTAATGCAATTTTCAACT 2598
1621 GGAATGATTAAGAAACTTTGAAATTTTACAGAGTGAAGAAATTAATGCAATTTTCAACT 1680
2599 TTTTACCTCTCCGAGAGCTAAATGAGAAAGAAAGATTTTCTTCTGAGGAGAGAGCTGTTTCCG 2658
1681 TTTTACCTCTCTCCGAGAGCTAAATGAGAAAGAAAGATTTTCTTCTGAGGAGAGAGCTGTTTCCG 1740
2659 TGGGTAGAACAGAGAAATTTATAGCAAGTATCCGTTGAAAAGAGCTAAATCTTGCA 2718
1741 TGGGTAGAACAGAGAAATTTATAGCAAGTATCCGTTGAAAAGAGCTAAATCTTGCA 1800
2719 AAAAAAGATTTTAAATATTAAGAACTATCAATTAAGCAAGAACTGAATTTTATGTTGAAAT 2778
1801 AAAAAAGATTTTAAATATTAAGAACTATCAATTAAGCAAGAACTGAATTTTATGTTGAAAT 1860
2779 ATTGGAATCTGCGAAGCAAGTATCCGGAAGTCAAGTCCGAGGAAGATTAATATGTTCAA 2838
1861 ATTGGAATCTGCGAAGCAAGTATCCGGAAGTCAAGTCCGAGGAAGATTAATATGTTCAA 1920
2839 AGATCGAAAAATTTAGCTATCTGATGAAGCTTAAAGAAAGCTGAATTTATCAGAGAAAT 2898
1921 AGATCGAAAAATTTAGCTATCTGATGAAGCTTAAAGAAAGCTGAATTTATCAGAGAAAT 1980
2899 ATTATGCAAGTGAATTTGAACAGACTTTTTCATGATGAGCGGATCTTTTATATGTTGCTCA 2958
1981 ATTATGCAAGTGAATTTGAACAGACTTTTTCATGATGAGCGGATCTTTTATATGTTGCTCA 2040
2959 GGTGGAATGCAATCAATTTGGAATGGAAGTATAGTGAAGTATCAGTAAGCGAAAGATT 3018
2041 GGTGGAATGCAATCAATTTGGAATGGAAGTATAGTGAAGTATCAGTAAGCGAAAGATT 2100
3019 TCCATTTGATGAGAGCAATTTTGAAGCTAATTAATAAATTTGCTTTTAAACAGTAAGAT 3078
2101 TCCATTTGATGAGAGCAATTTTGAAGCTAATTAATAAATTTGCTTTTAAACAGTAAGAT 2160
3079 GATACTTCTGTTTGAATGCTGCGGTTCAAGCGGAAATCGGAACGAAAAATGCGCGGCTC 3138

QY 8545 CAAGGAACAAACAAACATTTTACGAGATCAACAAAGTTAACTTCTACAACTGTAAATGGA 8604
DB 1141 CAAGGAACAAACAAACATTTTACGAGATCAACAAAGTTAACTTCTACAACTGTAAATGGA 1200
QY 8605 GGAACACGATCTCAACCTTGTGCAAAAGCTTGGCTTAAATAAGAAATTTATGGAATGTA 8664
DB 1201 GGAACACGATCTCAACCTTGTGCAAAAGCTTGGCTTAAATAAGAAATTTATGGAATGTA 1260
QY 8665 AAAGGAACGTGAGAGACCTTGTGAGAGCGAAACGACGCGCTTGAATAATTAACAAG 8724
DB 1261 AAAGGAACGTGAGAGACCTTGTGAGAGCGAAACGACGCGCTTGAATAATTAACAAG 1320
QY 8725 AGTACTACAGAGCAATGGTTGCAAGAAATTTGGGAAATTTGAGATTAATTAAGAAACGAT 8784
DB 1321 AGTACTACAGAGCAATGGTTGCAAGAAATTTGGGAAATTTGAGATTAATTAAGAAACGAT 1380
QY 8785 GCAAGAGATTAATCACTTTGTAAAGTCAACGAGACGAAACCAAGAGAGCTTTGTCCGA 8844
DB 1381 GCAAGAGATTAATCACTTTGTAAAGTCAACGAGACGAAACCAAGAGAGCTTTGTCCGA 1440
QY 8845 AAGATGGTATTTCTGTGAAATAACAATTTGAGGGGAAACAAATTCATCCATGGAAGT 8904
DB 1441 AAGATGGTATTTCTGTGAAATAACAATTTGAGGGGAAACAAATTCATCCATGGAAGT 1500
QY 8905 AAAGCCAGAAATTTGTGAAACCGGAAAGTGTAAATGTAGATGCTTTGAAATGAACCTTGATGA 8964
DB 1501 AAAGCCAGAAATTTGTGAAACCGGAAAGTGTAAATGTAGATGCTTTGAAATGAACCTTGATGA 1560
QY 8965 GATCTACAAAGGAAAGAGTGGCTATGTGTGAATTTGGTATTTGGAATTTGATGTAAT 9024
DB 1561 GATCTACAAAGGAAAGAGTGGCTATGTGTGAATTTGGTATTTGGAATTTGATGTAAT 1620
QY 9025 AATGTGATTAAGAAATAATGTAGAGCGAAATCGAAGACATGCTATTGTAGAACTACT 9084
DB 1621 AATGTGATTAAGAAATAATGTAGAGCGAAATCGAAGACATGCTATTGTAGAACTACT 1680
QY 9085 GGAACAAACAAAGATTCACAGCATTTCACAGACGAAAGTAATATTTCTTGAAAGAGAGC 9144
DB 1681 GGAACAAACAAAGATTCACAGCATTTCACAGACGAAAGTAATATTTCTTGAAAGAGAGC 1740
QY 9145 GGTGACAGCTGACGCTGCAATATCGAATGTACACATTTCCATGAGATGATTTAAAT 9204
DB 1741 GGTGACAGCTGACGCTGCAATATCGAATGTACACATTTCCATGAGATGATTTAAAT 1800
QY 9205 TTGGCAAGACGATGATGATCTTCTCAATTAATTAACCAAAATTTCAAAAATAATTAAT 9264
DB 1801 TTGGCAAGACGATGATGATCTTCTCAATTAATTAACCAAAATTTCAAAAATAATTAAT 1860
QY 9265 TTAGCATCAAGTATGATCGAATGTGAATGTTCAATGGGTGGCTGAAGCAAGAGTGA 9324
DB 1861 TTAGCATCAAGTATGATCGAATGTGAATGTTCAATGGGTGGCTGAAGCAAGAGTGA 1920
QY 9325 GGAACCAAGCAAGCTTGTGTAAAGATCAATTAATTAAGATTAATTTGATTTA 9384
DB 1921 GGAACCAAGCAAGCTTGTGTAAAGATCAATTAATTAAGATTAATTTGATTTA 1980
QY 9385 GCAAGGAAATTTAAACAGAGGAAACATCAATGTATATGCGGATATGATAAAATTAAT 9444
DB 1981 GCAAGGAAATTTAAACAGAGGAAACATCAATGTATATGCGGATATGATAAAATTAAT 2040
QY 9445 AATATTAAGTACAAATTTCTAAGGCTATTTGCGATGCGCAAAAGTCATGCTCAGCTGCT 9504
DB 2041 AATATTAAGTACAAATTTCTAAGGCTATTTGCGATGCGCAAAAGTCATGCTCAGCTGCT 2100
QY 9505 TTGGCACTGCGCACTATTGAAAAAATGAAGTAATTTAATTAATGCGATCCGAGAAATTT 9564
DB 2101 TTGGCACTGCGCACTATTGAAAAAATGAAGTAATTTAATTAATGCGATCCGAGAAATTT 2160
QY 9565 AAAAATTAATCTGCGCAAGATTGGAAGGGAAGGTAATTAATAAAACGTCGGTATGAGATCTAAT 9624
DB 2161 AAAAATTAATCTGCGCAAGATTGGAAGGGAAGGTAATTAATAAAACGTCGGTATGAGATCTAAT 2220

QY 9625 CAGGTAGACTGTATACGGATTAATATACATGCAATTTCTGTAAAAACATTAACAAAA 9684
DB 2221 CAGGTAGACTGTATACGGATTAATATACATGCAATTTCTGTAAAAACATTAACAAAA 2280
QY 9685 TTGACATTAATCAATCAAGAGAGGAGAAAAAGGAAAAATTA 9726
DB 2281 TTGACATTAATCAATCAAGAGAGGAGAAAAAGGAAAAATTA 2322

RESULT 8
US-10-647-057-13
; Sequence 13, Application US/10647057
; Publication No. US20040047871A1
; GENERAL INFORMATION:
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION
; APPLICANT: NAGARAJA, T.
; APPLICANT: STEWART, GEORGE
; APPLICANT: MARYANAN, SANJEEV
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP
; FILE REFERENCE: PCT-30962
; CURRENT APPLICATION NUMBER: US/10/647,057
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 2322
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-10-647-057-13

Query Match 23.9%; Score 2322; DB 7; Length 2322;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7405 GAAATTTAAAGCCATTGTGAGAGCTTCTGATATTAATTTGAAAAAATGTAGATATTACACA 7464
DB 1 GAAATTTAAAGCCATTGTGAGAGCTTCTGATATTAATTTGAAAAAATGTAGATATTACACA 60
QY 7465 GAGACAAAGATTAATATCACTTCTACTGCTGTTGGAACTGACAGCTTCTGCTCCGA 7524
DB 61 GAGACAAAGATTAATATCACTTCTACTGCTGTTGGAACTGACAGCTTCTGCTCCGA 120
QY 7525 TCAGGAACAGTGGCAGTTACAAATATTTAAAGAAATTTCCGAGTTACTGTGAAATTTCT 7584
DB 121 TCAGGAACAGTGGCAGTTACAAATATTTAAAGAAATTTCCGAGTTACTGTGAAATTTCT 180
QY 7585 TTTGTGAAAGCAGCTGAAAAAGTAATGTTAGATCGATATTACAGAAATGTTGCTTTA 7644
DB 181 TTTGTGAAAGCAGCTGAAAAAGTAATGTTAGATCGATATTACAGAAATGTTGCTTTA 240
QY 7645 ACAGCATATCAAGTCCCTGTAGAGACATTTGGGAAATGAGAGCTGCTATGCAATTAAT 7704
DB 241 ACAGCATATCAAGTCCCTGTAGAGACATTTGGGAAATGAGAGCTGCTATGCAATTAAT 300
QY 7705 TCTAATGGAAGATCAAAATATCAATTAATTTAAATTTCTAAGTATTAGAAAAAATTTGAT 7764
DB 301 TCTAATGGAAGATCAAAATATCAATTAATTTAAATTTCTAAGTATTAGAAAAAATTTGAT 360
QY 7765 GTTATTGTAAGATTAATTCGAAATTTGAGACGGAACCAAAAGATTAAACGATAGAGCG 7824
DB 361 GTTATTGTAAGATTAATTCGAAATTTGAGACGGAACCAAAAGATTAAACGATAGAGCG 420
QY 7825 GTAGTCCGAGAGCATTATCTCAAAAGCAAGATGAATGAATTCAGAGGTTGAAATTT 7884
DB 421 GTAGTCCGAGAGCATTATCTCAAAAGCAAGATGAATGAATTCAGAGGTTGAAATTT 480
QY 7885 GAGAAAGTATTTTCAATGAAGAAATGAGATTAATGAGAGCTTCTTAAAGAAATTTGAAAGA 7944
DB 481 GAGAAAGTATTTTCAATGAAGAAATGAGATTAATGAGAGCTTCTTAAAGAAATTTGAAAGA 540
QY 7945 GAAATCAATGTCAAGTGGAAAAAGAAAAAGAGTGAATGCTGTAATCTCAAGAGCTTCT 8004

Db 541 GAAATCAATGTCACAGTGGAAAAAGAAAAAGAGTACCTGTAATCTCCAGACCTTCT 600
Qy 8005 GTAGAGACAGTACGAGGGGAGAAATTTTCCGAGCAAAAGATCCGGAAGCTCTTAT 8064
Db 601 GTAGAGACAGTACGAGGGGAGAAATTTTCCGAGCAAAAGATCCGGAAGCTCTTAT 660
Qy 8065 TTGAAAGTTAGTACAAAATCCGGAGAAAGATTTTTCATGCAGATATGTGAATGTGAA 8124
Db 661 TTGAAAGTTAGTACAAAATCCGGAGAAAGATTTTTCATGCAGATATGTGAATGTGAA 720
Qy 8125 GCACACATTAATAAAGTAAACAGCTTTCTAAGAGTAAAGAGTAAAGGTTCTGTATGGA 8184
Db 721 GCACACATTAATAAAGTAAACAGCTTTCTAAGAGTAAAGAGTAAAGGTTCTGTATGGA 780
Qy 8185 GGAGTTGAGTACCAAGGAGCAAGCTACTGCTGAGGTAATACTATGATAGAGTTGAG 8244
Db 781 GGAGTTGAGTACCAAGGAGCAAGCTACTGCTGAGGTAATACTATGATAGAGTTGAG 840
Qy 8245 GAAGGAAATTTGTCAGAACAATGATGAAATGCAATTTCTAAGTAAAGTGGTTGGAT 8304
Db 841 GAAGGAAATTTGTCAGAACAATGATGAAATGCAATTTCTAAGTAAAGTGGTTGGAT 900
Qy 8305 GAAGTAAAGTAACTGCTTAAATCTTCTGATGATACAGAAATGAGAGGAAATGCCGGA 8364
Db 901 GAAGTAAAGTAACTGCTTAAATCTTCTGATGATACAGAAATGAGAGGAAATGCCGGA 960
Qy 8365 GAGAGAGTAACTGCTTAAATCTTCTGATGATACAGAAATGAGAGGAAATGCCGGA 8424
Db 961 GAGAGAGTAACTGCTTAAATCTTCTGATGATACAGAAATGAGAGGAAATGCCGGA 1020
Qy 8425 CAAGATTAAGAAATTAATGATTAACAACAAATAATATATTTTCAAGTCAATCTCTTGT 8484
Db 1021 CAAGATTAAGAAATTAATGATTAACAACAAATAATATATTTTCAAGTCAATCTCTTGT 1080
Qy 8485 TTTAATGATACAAAGATGAGAGATATGAAATCTTTAGCCGTAGCCGTGCGATGGA 8544
Db 1081 TTTAATGATACAAAGATGAGAGATATGAAATCTTTAGCCGTAGCCGTGCGATGGA 1140
Qy 8545 CAAGGAAACAAACAAAGCTTTTACAGATCAACAAAGTAACTTCTCAACTGTAAATGGA 8604
Db 1141 CAAGGAAACAAACAAAGCTTTTACAGATCAACAAAGTAACTTCTCAACTGTAAATGGA 1200
Qy 8605 GGAAGATGATCTCAACTGCTGCAAAAGCTTTGGCTTAAATGAAATATATGAAATGTA 8664
Db 1201 GGAAGATGATCTCAACTGCTGCAAAAGCTTTGGCTTAAATGAAATATATGAAATGTA 1260
Qy 8665 AAAGGAACTGAGAGAGCTTTAGTCCGAGCGGAAACGACGCCGTGAAATTTTACAAG 8724
Db 1261 AAAGGAACTGAGAGAGCTTTAGTCCGAGCGGAAACGACGCCGTGAAATTTTACAAG 1320
Qy 8725 AGTACTACAGAGAGCTTTGCTGAGAAATTTGGAAATTTGAGATTAATTTAGAAAGAT 8784
Db 1321 AGTACTACAGAGAGCTTTGCTGAGAAATTTGGAAATTTGAGATTAATTTAGAAAGAT 1380
Qy 8785 GCAGAGATTAATGATGATGAGTCAACGAGAGCGAAACCAAGAGAGCTTTGTCGA 8844
Db 1381 GCAGAGATTAATGATGATGAGTCAACGAGAGCGAAACCAAGAGAGCTTTGTCGA 1440
Qy 8845 AAGATGATGATTTCTGTAAGAAATTAATCAATTTCCGGGAAACAAATTCATTCATTAAGAT 8904
Db 1441 AAGATGATGATTTCTGTAAGAAATTAATCAATTTCCGGGAAACAAATTCATTCATTAAGAT 1500
Qy 8905 AAAGCCAGATTTGTCGAACCGGAGATTAATGATGATCTTTGAATGAATTTGATGTA 8964
Db 1501 AAAGCCAGATTTGTCGAACCGGAGATTAATGATGATCTTTGAATGAATTTGATGTA 1560
Qy 8965 GATCTACAGAGAAAGTGTGCTATGTGAAATTTGTAATTTGAAATTTGATGTAAT 9024
Db 1561 GATCTACAGAGAAAGTGTGCTATGTGAAATTTGTAATTTGTAATTTGATGTAAT 1620
Qy 9025 AATGTGATTAAGAAATTTGAGAGCCAAATCGGAAGACATCTATTTGTAAGAACTACT 9084

Db 1621 AATGTGATTAAGAAATTTGAGAGCCAAATCCGAGACATGCTATTTGAGAACTACT 1680
Qy 9085 GGAAGAAACAAATATCAAGCTTTTACAAGACAAAGTAAATTTCTTGGAAAAAGAGAC 9144
Db 1681 GGAAGAAACAAATATCAAGCTTTTACAAGACAAAGTAAATTTCTTGGAAAAAGAGAC 1740
Qy 9145 GCTGAGCTGAGCTGCAATATCGAATGTACATTTTCCAAATGAGATGATTTAAAAAT 9204
Db 1741 GCTGAGCTGAGCTGCAATATCGAATGTACATTTTCCAAATGAGATGATTTAAAAAT 1800
Qy 9205 TTGGCAAGCAGTATGATCTTCTCAATTAATTAACCAAAATTCAAAAATTAATTTACT 9264
Db 1801 TTGGCAAGCAGTATGATCTTCTCAATTAATTAACCAAAATTCAAAAATTAATTTACT 1860
Qy 9265 TTAGCATCAATGATGAAATCGAATGTGAATTTTCAATGGGGTGGCTGAAGCAAGAGTGA 9324
Db 1861 TTAGCATCAATGATGAAATCGAATGTGAATTTTCAATGGGGTGGCTGAAGCAAGAGTGA 1920
Qy 9325 GAGGCCAAAGGACAGTATGATTAAGATCAATTAATTAAGTAAATGATTTGATTTTA 9384
Db 1921 GAGGCCAAAGGACAGTATGATTAAGATCAATTAATTAAGTAAATGATTTGATTTTA 1980
Qy 9385 GCAGGAAATTAATTAACAGAGGAAACATCAATGATATGCCGATATGATTAATTAAT 9444
Db 1981 GCAGGAAATTAATTAACAGAGGAAACATCAATGATATGCCGATATGATTAATTAAT 2040
Qy 9445 AATATTAAGTAAACAAATTTTCAAGCTTTTGGGATGCGCAAAAGTCACTGCTGCT 9504
Db 2041 AATATTAAGTAAACAAATTTTCAAGCTTTTGGGATGCGCAAAAGTCACTGCTGCT 2100
Qy 9505 TCGGCACTGCCATTTGAAAAAAATGAAATTAATTAATTAATGCGATCCGAGATTT 9564
Db 2101 TCGGCACTGCCATTTGAAAAAAATGAAATTAATTAATTAATGCGATCCGAGATTT 2160
Qy 9565 AAAAAATTAATCTGCAAGATTTGAGAGGAAAGCTAAATAAAAAGCTGCTAGATCTAAT 9624
Db 2161 AAAAAATTAATCTGCAAGATTTGAGAGGAAAGCTAAATAAAAAGCTGCTAGATCTAAT 2220
Qy 9625 CAGGTAGCTGATTAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9684
Db 2221 CAGGTAGCTGATTAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
Qy 9685 TTGACATATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9726
Db 2281 TTGACATATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2322

RESULT 9
US-09-841-786-11
; Sequence 11, Application US/09841786
; Patent No. US2002005483A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; PRIOR APPLICATION NUMBER: 2001-04-24
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2141
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-11

Query Match 22.0%; Score 2141; DB 3; Length 2141;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3553 GGAGCTTCTGAAAGTGTTCCTTCAATAATATTAAGAAACATCTGCTGTAGAT 3612
Db 1 GGAGCTTCTGAAAGTGTTCCTTCAATAATATTAAGAAACATCTGCTGTAGAT 60
OY 3613 GGAGTAAAGATTAATTTGAAGGAGCAATAAAGTAGAGTGACTTCTGTATCT 3672
Db 61 GGAGTAAAGATTAATTTGAAGGAGCAATAAAGTAGAGTGACTTCTGTATCT 120
OY 3673 ACTTTTGTGAGCATGGGGCGGATCTGTGCATCTTCACTGATGATCATTTGGAAGTGA 3732
Db 121 ACTTTTGTGAGCATGGGGCGGATCTGTGCATCTTCACTGATGATCATTTGGAAGTGA 180
OY 3723 AATAGCAACATCAGTGTGCTTGTAGCTGAGGGGCTGTAAATTAATTTCAAGTAA 3792
Db 181 AATAGCAACATCAGTGTGCTTGTAGCTGAGGGGCTGTAAATTAATTTCAAGTAA 240
OY 3793 ACAAGTCTTTGGTTAAAAATAGTATATTCGAAATGCCAATAAATTTAAATATGCT 3852
Db 241 ACAAGTCTTTGGTTAAAAATAGTATATTCGAAATGCCAATAAATTTAAATATGCT 300
OY 3853 TTGAGTGAAGAACTCAAGTACAGAGAGAGAGAGGTTTGAAGCACTTAAAGAAAGTGA 3912
Db 301 TTGAGTGAAGAACTCAAGTACAGAGAGAGAGAGGTTTGAAGCACTTAAAGAAAGTGA 360
OY 3913 GGACAGAGAAAAAGTTATCTATTTGGGAACTTCTGCTTCTATCATCTTGTGTGAACAATGA 3972
Db 361 GGACAGAGAAAAAGTTATCTATTTGGGAACTTCTGCTTCTATCATCTTGTGTGAACAATGA 420
OY 3973 GTTCTGCAAAATCAGAAATTAATAAGTAGAGAGAGAACTGAAAGCCAAATAATGAT 4032
Db 421 GTTCTGCAAAATCAGAAATTAATAAGTAGAGAGAGAACTGAAAGCCAAATAATGAT 480
OY 4033 GTTGAATGCACCTGCTTATCAAGCGGACACCCAGTGAAGAGAGCTTTAAATTTACAAGCT 4092
Db 481 GTTGAATGCACCTGCTTATCAAGCGGACACCCAGTGAAGAGAGCTTTAAATTTACAAGCT 540
OY 4093 GGAAGAGTCAAAATGAGTGTAGGGGCTACTGTGACTGTGGCCAAATTTAAACAACAAGTA 4152
Db 541 GGAAGAGTCAAAATGAGTGTAGGGGCTACTGTGACTGTGGCCAAATTTAAACAACAAGTA 600
OY 4153 AATGCTTCTATTAGTGTGGAGATATACCTAACTGAAATCGAGCGAGCAAAAGCTTT 4212
Db 601 AATGCTTCTATTAGTGTGGAGATATACCTAACTGAAATCGAGCGAGCAAAAGCTTT 660
OY 4213 TTAGAACAACCTCAAGTGACTGCTGCAAGTGAAGAGAGGAGCAATTAATTTCTGAGCG 4272
Db 661 TTAGAACAACCTCAAGTGACTGCTGCAAGTGAAGAGAGGAGCAATTAATTTCTGAGCG 720
OY 4273 GGAATTAGAGAAATTTCAAGGGGCTGTTTCTGTCAATTAAGATTGACAATGAGCTGAGCT 4332
Db 721 GGAATTAGAGAAATTTCAAGGGGCTGTTTCTGTCAATTAAGATTGACAATGAGCTGAGCT 780
OY 4333 AGCGTTGATAAATCTTCACTGAGAGAGCTAAATGAATCATGTCATTTGCCAAAGATGTC 4392
Db 781 AGCGTTGATAAATCTTCACTGAGAGAGCTAAATGAATCATGTCATTTGCCAAAGATGTC 840
OY 4393 AAAAGAACTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGAGAAAGATTAAGAA 4452
Db 841 AAAAGAACTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGAGAAAGATTAAGAA 900
OY 4453 TATTAGAAAGATCGTGTATTAAATGAGCTGAGAAATGGTTATTATAGAGAGAAACAATA 4512
Db 901 TATTAGAAAGATCGTGTATTAAATGAGCTGAGAAATGGTTATTATAGAGAGAAACAATA 960
OY 4513 GAAAAAGCAAAAGAAAAAGAGAGCGGTCAATTTGTAATGCTGCTTTATCGGTTGCTGA 4572
Db 961 GAAAAAGCAAAAGAAAAAGAGAGCGGTCAATTTGTAATGCTGCTTTATCGGTTGCTGA 1020
OY 4573 ACGGATTAATTCGCTGAGAGAGTACATATTTGCAATGCTTAATTTAAATTTAA 4632
Db 1021 ACGGATTAATTCGCTGAGAGAGTACATATTTGCAATGCTTAATTTAAATTTAA 1080

OY 4633 GCAGAAATTGAGTGAAGCAATTAAGGAGCGGAGAGATTAATTAATTCATGCGAAACATGTA 4692
Db 1081 GCAGAAATTGAGTGAAGCAATTAAGGAGCGGAGAGATTAATTAATTCATGCGAAACATGTA 1140
OY 4693 AATGAGAGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGCACTTGCTATCAGC 4752
Db 1141 AATGAGAGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGCACTTGCTATCAGC 1200
OY 4753 AAAAGTCTTTTTCAGGAAATGGGATCTGAGACATGGCAAGACTTATCAATGACAGATT 4812
Db 1201 AAAAGTCTTTTTCAGGAAATGGGATCTGAGACATGGCAAGACTTATCAATGACAGATT 1260
OY 4813 GCAAGGTGAGTAAAGGAAAGATTTCTGCTGATTCCTTAAATGTAAGCAAAATTAATCC 4872
Db 1261 GCAAGGTGAGTAAAGGAAAGATTTCTGCTGATTCCTTAAATGTAAGCAAAATTAATCC 1320
OY 4873 ATTCTTGGGGTGAATGTTTGGCGGAAACAATGCCGTTCTCTTCTACCGCGGTAGAGCT 4932
Db 1321 ATTCTTGGGGTGAATGTTTGGCGGAAACAATGCCGTTCTCTTCTACCGCGGTAGAGCT 1380
OY 4933 GCTTTTGGCAATTAATCTCTTCTATTAATTAATCTCTGCTTTGATTAACGAAACGAAGTA 4992
Db 1381 GCTTTTGGCAATTAATCTCTTCTATTAATTAATCTCTGCTTTGATTAACGAAACGAAGTA 1440
OY 4993 AATCCTTTTGTGGAAGAAATACAAAGTCAATGATCAAGGCTTGAATGATTTCTATAT 5052
Db 1441 AATCCTTTTGTGGAAGAAATACAAAGTCAATGATCAAGGCTTGAATGATTTCTATAT 1500
OY 5053 ACAAAAGTTTCTGCTGAGAGCGCTGCAAGTATTAAGAGGCTGGAATGGAAGATGTA 5112
Db 1501 ACAAAAGTTTCTGCTGAGAGCGCTGCAAGTATTAAGAGGCTGGAATGGAAGATGTA 1560
OY 5113 TCTGTCAATCGTGTCTGATGAAACGGAACCTTAAGTATGATTCGAATTTGAAGA 5172
Db 1561 TCTGTCAATCGTGTCTGATGAAACGGAACCTTAAGTATGATTCGAATTTGAAGA 1620
OY 5173 GTAAGTCTTTCATATGATGACAAAGATCAAAAAACAATTAATTAATTAATGCGGGAAT 5232
Db 1621 GTAAGTCTTTCATATGATGACAAAGATCAAAAAACAATTAATTAATTAATGCGGGAAT 1680
OY 5233 GCAATGAGAGAAAAAGCGGCTGAGGTTGAGCAACAGTGTCTCATCAATATTTGAAAA 5292
Db 1681 GCAATGAGAGAAAAAGCGGCTGAGGTTGAGCAACAGTGTCTCATCAATATTTGAAAA 1740
OY 5293 CAATCAGTTATAGCTATTTGTAAGAAACAGTAAATTTACAGCGGGAATGATCAAGATGA 5352
Db 1741 CAATCAGTTATAGCTATTTGTAAGAAACAGTAAATTTACAGCGGGAATGATCAAGATGA 1800
OY 5353 AAAAATATCAATGATGACGCAAAAGATTAATCAATGACCAATTAATTAATGAGAGGAT 5412
Db 1801 AAAAATATCAATGATGACGCAAAAGATTAATCAATGACCAATTAATTAATGAGAGGAT 1860
OY 5413 GAGAGAGCAAAAGAGGCTCTGTGCAAGAGCTTCTGCAAGTACTACCTTGAATTAAGCA 5472
Db 1861 GAGAGAGCAAAAGAGGCTCTGTGCAAGAGCTTCTGCAAGTACTACCTTGAATTAAGCA 1920
OY 5473 GTTCTTCTCATGTTGATCAAACTGATTTGAACAAAGTTTGAAGAAAGAAATTAATGA 5532
Db 1921 GTTCTTCTCATGTTGATCAAACTGATTTGAACAAAGTTTGAAGAAAGAAATTAATGA 1980
OY 5533 AATAAGGAAAGGCAAAATGTTAATGTTCTAGCTGAAAAATAGAGTCAAGGTGTCACAAAT 5592
Db 1981 AATAAGGAAAGGCAAAATGTTAATGTTCTAGCTGAAAAATAGAGTCAAGGTGTCACAAAT 2040
OY 5593 GCGACAGTGTCTTCCGAGCAAGTGAACAAGCTCAGTAGAGAGCTGAGTAGAGATTAAAT 5652
Db 2041 GCGACAGTGTCTTCCGAGCAAGTGAACAAGCTCAGTAGAGAGCTGAGTAGAGATTAAAT 2100
OY 5653 AAAATTTACAAATTAATCTTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 5693
Db 2101 AAAATTTACAAATTAATCTTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2141

```
RESULT 10
US-10-647-057-11
; Sequence 11, Application US/10647057
; Publication No. US20040047871A1
; GENERAL INFORMATION:
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION
; APPLICANT: NAGARAJA, T.
; APPLICANT: STEWART, GEORGE
; APPLICANT: NARAYANAN, SANJEEV
; APPLICANT: CHENGAPPA, M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP
; TITLE OF INVENTION: THEROF
; FILE REFERENCE: PCT-30962
; CURRENT APPLICATION NUMBER: US/10/647,057
; CURRENT FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 2141
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-10-647-057-11

Query Match      22.0%; Score 2141; DB 7; Length 2141;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3553 GGAGGCTTCGGAAGGTTCTTTCAATATATTTAAAGAAACATCTGCTGCTAGAT 3612
DB 1 GGAGCTTCGGAAGGTTCTTTCAATATATTTAAAGAAACATCTGCTGCTAGAT 60
QY 3613 GGAGTAAGATTAATTTGAAGGAGCAATTAAGAGTAGAGTCTTTCTGATTC 3672
DB 61 GGAGTAAGATTAATTTGAAGGAGCAATTAAGAGTAGAGTCTTTCTGATTC 120
QY 3673 ACTTTTGTGGAGCATGGGCGGATCTGCTGCACTTCAGTGAATATTTGGAAGTGA 3732
DB 121 ACTTTTGTGGAGCATGGGCGGATCTGCTGCACTTCAGTGAATATTTGGAAGTGA 180
QY 3733 AATAGCAATAGTCTGTTAGCTGAGGCGGCTGTAAATTAATTTCAAGTAA 3792
DB 181 AATAGCAATAGTCTGTTAGCTGAGGCGGCTGTAAATTAATTTCAAGTAA 240
QY 3793 ACAAGTCTTGGTTAAATAAGTAGATATTCGAATGCCAATTAATTTAAAGTAA 3852
DB 241 ACAAGTCTTGGTTAAATAAGTAGATATTCGAATGCCAATTAATTTAAAGTAA 300
QY 3853 TTGAGTGAAGAACTCAAGTAGACAGACAGAGGTTTGGAAAGCACTTAAAGTGA 3912
DB 301 TTGAGTGAAGAACTCAAGTAGACAGACAGAGGTTTGGAAAGCACTTAAAGTGA 360
QY 3913 GGAAGAGGAAAGATATCATCTATGGGAACCTTCGCTTCATCAACTTAGAACAATGA 3972
DB 361 GGAAGAGGAAAGATATCATCTATGGGAACCTTCGCTTCATCAACTTAGAACAATGA 420
QY 3973 GTTTCGCAAAATCAGAAAAATATACAGTAGCAGAGAAATCTGAAAGCCAAAAATGAT 4032
DB 421 GTTTCGCAAAATCAGAAAAATATACAGTAGCAGAGAAATCTGAAAGCCAAAAATGAT 480
QY 4033 GTTGATGTCACTGCTTATCAAGCGGACACCCAAAGTGAACAGAGCTTTAAATTTCAAGCT 4092
DB 481 GTTGATGTCACTGCTTATCAAGCGGACACCCAAAGTGAACAGAGCTTTAAATTTCAAGCT 540
QY 4093 GGAAGTCAAAATGGAATCTGAGGGGCTACTGTGACTGTTGCCAAATTTAAACAAGTA 4152
DB 541 GGAAGTCAAAATGGAATCTGAGGGGCTACTGTGACTGTTGCCAAATTTAAACAAGTA 600
QY 4153 AATGCTTCTAATTAGTGAGGAGATATCTAATGTTAATCGACGGAACGCAAAAGCTCTT 4212
DB 601 AATGCTTCTAATTAGTGAGGAGATATCTAATGTTAATCGACGGAACGCAAAAGCTCTT 660
QY 4213 TTGCAACCACTCAAGTGAAGTCTGCAAGTGAACGAGGAGCAATTAAGTTCTGAGCG 4272
```

```
DB 661 TTGCAACCACTCAAGTGAAGTCTGCAAGTGAACGAGGAGCAATTAAGTTCTGAGCG 720
QY 4273 GGATTTAGAAATTTATCAAGGGGCTGTTCTGTCAATTAAGTTGACATGACGTGAAGCT 4332
DB 721 GGATTTAGAAATTTATCAAGGGGCTGTTCTGTCAATTAAGTTGACATGACGTGAAGCT 780
QY 4333 AGCGTTGATAAATCTTCATGGAAGGAGCTAATGAATCAATGTCATTCAGCAAGATGTC 4392
DB 781 AGCGTTGATAAATCTTCATGGAAGGAGCTAATGAATCAATGTCATTCAGCAAGATGTC 840
QY 4393 AAAGAGTTTGTGATCTAGCAAAAGATATATGAGCTTTACTAATGAAAAATTAATAA 4452
DB 841 AAAGAGTTTGTGATCTAGCAAAAGATATATGAGCTTTACTAATGAAAAATTAATAA 900
QY 4453 TATTTAGAGATCGTGATTAATTAATGACGTGAATGTTATATGAGAGACACTA 4512
DB 901 TATTTAGAGATCGTGATTAATTAATGACGTGAATGTTATATGAGAGACACTA 960
QY 4513 GAAAAAGCAAAAGAAAAAGAGCGGCTCATTTGTAATGCTTTAATCGGTTGCTGA 4572
DB 961 GAAAAAGCAAAAGAAAAAGAGCGGCTCATTTGTAATGCTTTAATCGGTTGCTGA 1020
QY 4573 ACGGTAATATCCGCTGAGAGTAGCTATTGCAATCAATCTGTTAAATTAATTTAA 4632
DB 1021 ACGGTAATATCCGCTGAGAGTAGCTATTGCAATCAATCTGTTAAATTAATTTAA 1080
QY 4633 GCAGATTAAGTGAAGCAATTAAGAGCCGAGAGATTAATTAATCATGCGAAACATGTA 4692
DB 1081 GCAGATTAAGTGAAGCAATTAAGAGCCGAGAGATTAATTAATCATGCGAAACATGTA 1140
QY 4693 AATGTGAGGCAAAATCTATCTGTGTGTGGAATGGGGCTTCGACCTGCTATCAGC 4752
DB 1141 AATGTGAGGCAAAATCTATCTGTGTGTGGAATGGGGCTTCGACCTGCTATCAGC 1200
QY 4753 AAAAGTCTTTTTCAGAAATGGATCTGAGCATGGAAGCTTAATCAATGACAGAT 4812
DB 1201 AAAAGTCTTTTTCAGAAATGGATCTGAGCATGGAAGCTTAATCAATGACAGAT 1260
QY 4813 GCAGAGTGAATTAAGAGAAATTTCTGCTGATCTTTAATGTAAGCAATTAATTC 4872
DB 1261 GCAGAGTGAATTAAGAGAAATTTCTGCTGATCTTTAATGTAAGCAATTAATTC 1320
QY 4873 ATTCTTGGGATTAATGTTGCGGAAACATTCGCGGTTCTCTTTCAACGCGGTAGAGCT 4932
DB 1321 ATTCTTGGGATTAATGTTGCGGAAACATTCGCGGTTCTCTTTCAACGCGGTAGAGCT 1380
QY 4933 GCTTTTGGAAATTAATCTTCAATTAATAAATCTGCTTTGATTAACAGAAAGATGA 4992
DB 1381 GCTTTTGGAAATTAATCTTCAATTAATAAATCTGCTTTGATTAACAGAAAGATGA 1440
QY 4993 AATCTTTTATGTAAGAAATACAAAGTCAATGTAACAGCTTTGAATGATTTCTATAT 5052
DB 1441 AATCTTTTATGTAAGAAATACAAAGTCAATGTAACAGCTTTGAATGATTTCTATAT 1500
QY 5053 ACAAAAGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGGAATGGAAGATGTA 5112
DB 1501 ACAAAAGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGGAATGGAAGATGTA 1560
QY 5113 TCTGTCAATCGGTGTTGATGAAACGGAAGCTTTAGTAGATTTCTGATTTGAAGA 5172
DB 1561 TCTGTCAATCGGTGTTGATGAAACGGAAGCTTTAGTAGATTTCTGATTTGAAGA 1620
QY 5173 GTAAGTTCTTTCAATGATGATCAAAAGATCAAAAGATCAATTAATCAATTCGCGAAAT 5232
DB 1621 GTAAGTTCTTTCAATGATGATCAAAAGATCAAAAGATCAATTAATCAATTCGCGAAAT 1680
QY 5233 GCAATGAGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACAAATTTTGAAGA 5292
DB 1681 GCAATGAGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACAAATTTTGAAGA 1740
QY 5293 CAATCAGTATAGCTATTGTAATAAACAATTAATTAACAGCGCAATGATCAAGTAGA 5352
```


QY 7004 GAGAGCCGCGAGCTTGTGACAGATTACCGAGTGGTTTCTGTAAACAATAAATATT 7063
DB 1381 GAGGAGCCGCGAGCTTGTGACAGATTACCGAGTGGTTTCTGTAAACAATAAATATT 1440
QY 7064 CTGTATAGCTGAGTTGACAAATACTGATTTGACTTCGGTAGAGAAAAGTAATG 7123
DB 1441 CTGTATAGCTGAGTTGACAAATACTGATTTGACTTCGGTAGAGAAAAGTAATG 1500
QY 7124 TAACGCAAAAGAGAAAATAATTTAAGCAAAACAGCAAAATCAGGAATCGAGGAG 7183
DB 1501 TAACGCAAAAGAGAAAATAATTTAAGCAAAACAGCAAAATCAGGAATCGAGGAG 1560
QY 7184 CAGCAATCGGAGCCAAATGCTTGTAAATATTTTGAACAGCTGTAGAAAGTAGAAAA 7243
DB 1561 CAGCAATCGGAGCCAAATGCTTGTAAATATTTTGAACAGCTGTAGAAAGTAGAAAA 1620
QY 7244 ATTCTGAAGAAAGAAAGAGAAAGTTTAAAACTTTAGACGAAGTTAACAAAGAACAG 7303
DB 1621 ATTCTGAAGAAAGAGAAAGAGTTTAAAACTTTAGACGAAGTTAACAAAGAACAG 1680
QY 7304 ATAAAAAGTAATGATGCTACGAAAAAAATCTTCAATCAGCAGGTAATTTCTACAGAG 7363
DB 1681 ATAAAAAGTAATGATGCTACGAAAAAAATCTTCAATCAGCAGGTAATTTCTACAGAG 1740
QY 7364 ATACTTCTTAAAGCGGATAGAGGATCTCAGGAGAGAGAAATTAAGCCATTGTA 7423
DB 1741 ATACTTCTTAAAGCGGATAGAGGATCTCAGGAGAGAGAAATTAAGCCATTGTA 1800
QY 7424 AGACTTCTGATTTATTTGAAAAATGTGATTTATCAACAGAGCAAGAAATTAATCA 7483
DB 1801 AGACTTCTGATTTATTTGAAAAATGTGATTTATCAACAGAGCAAGAAATTAATCA 1860
QY 7484 CTTCTACTGCTGTTTGGAGACTGCAG 7510
DB 1861 CTTCTACTGCTGTTTGGAGACTGCAG 1887

RESULT 12

US-10-647-057-12

; Sequence 12, Application US/10647057

; Publication No. US20040047871A1

; GENERAL INFORMATION:

; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION

; APPLICANT: NAGARAJA, T.

; APPLICANT: STEWART, GEORGE

; APPLICANT: NARAYANAN, SANJEEV

; APPLICANT: CHENGAPPA, M.

; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP

; FILE REFERENCE: PCT-30962

; CURRENT APPLICATION NUMBER: US/10/647,057

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 12

; LENGTH: 1887

; TYPE: DNA

; ORGANISM: Fusobacterium necrophorum

US-10-647-057-12

Query Match 19.4%; Score 1887; DB 7; Length 1887;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5624 CTGCAATGAGAGCTGAGTAGTAAATAATTAACAATAATCTGTCACATATA 5683
DB 1 CTGCAATGAGAGCTGAGTAGTAAATAATTAACAATAATCTGTCACATATA 60
QY 5684 AAAATAGTACTCAAAATGTACGAAATGCTTGTAAAGCAATCTCATTTATTA 5743
DB 61 AAAATAGTACTCAAAATGTACGAAATGCTTGTGTAAAGCAATCTCATTTATTA 120

QY 5744 AACCAATTGGAATTGAGAGCTGAGTAGGAGCTGAGAGAGCTGAGTAGACAGGTTCTGTAG 5803
DB 121 AAACAATTGGAATTGAGAGCTGAGTAGGAGCTGAGAGAGCTGAGTAGACAGGTTCTGTAG 180
QY 5804 CAGTGAATAGATTGTAAATATATGATAGAGAAATTAATCATCAAAAATCACTGCGA 5863
DB 181 CAGTGAATAGATTGTAAATATATGATAGAGAAATTAATCATCAAAAATCACTGCGA 240
QY 5864 AGGGAATGTGGAGTTTATCAGAGTGTATGCGGTAATTTGCTTAATTAATGAGAAAG 5923
DB 241 AGGGAATGTGGAGTTTATCAGAGTGTATGCGGTAATTTGCTTAATTAATGAGAAAG 300
QY 5924 TGTCTGAGTGGCCCGTGCAGCAATAGGAGCCTCAACAGTGTGAATGAAATTAACAGAT 5983
DB 301 TGTCTGAGTGGCCCGTGCAGCAATAGGAGCCTCAACAGTGTGAATGAAATTAACAGAT 360
QY 5984 CTACAAAAGCATATGTAAAGAAATTTCTACAGTGAATTTGCTAAAGAAAGAA 6043
DB 361 CTACAAAAGCATATGTAAAGAAATTTCTACAGTGAATTTGCTAAAGAAAGAA 420
QY 6044 TTACTACTCAAGGCGAAGTAGATTAAGTGTGATTAAGTATTCAAAAATCTTAATTA 6103
DB 421 TTACTACTCAAGGCGAAGTAGATTAAGTGTGATTAAGTATTCAAAAATCTTAATTA 480
QY 6104 AGGAAGCTTATCACAAAAAAGAAAAATAGTAAATTAAGAAAGATTGTTACCAATAGTT 6163
DB 481 AGGAAGCTTATCACAAAAAAGAAAAATAGTAAATTAAGAAAGATTGTTACCAATAGTT 540
QY 6164 CAGCTACTCATATCTTTAAATCTTTATTTGGCAATGCGCTGTGTTACAGACAGCCGAG 6223
DB 541 CAGCTACTCATATCTTTAAATCTTTATTTGGCAATGCGCTGTGTTACAGACAGCCGAG 600
QY 6224 TGGCAGAACTGTAAATATCAACAAAGTTTATGAGAAACAGAACTCTTGTAGAAATTT 6283
DB 601 TGGCAGAACTGTAAATATCAACAAAGTTTATGAGAAACAGAACTCTTGTAGAAATTT 660
QY 6284 CTATATTAATGCAAAACATTAATCTGTAAATCAGAGATTAACAGAAATTAATCGAG 6343
DB 661 CTATATTAATGCAAAACATTAATCTGTAAATCAGAGATTAACAGAAATTAATCGAG 720
QY 6344 TAGTAGTTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 6403
DB 721 TAGTAGTTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 780
QY 6404 TTATATAAAGAAATACCAAGACAGAGTTGCAAAATCTCAATGCTGATGAAGTTTGG 6463
DB 781 TTATATAAAGAAATACCAAGACAGAGTTGCAAAATCTCAATGCTGATGAAGTTTGG 840
QY 6464 GAGAAAGAGCTGAATTAACAGAGATTTCTAAGCAAGAAATTTCTTTTGGAGTCGAG 6523
DB 841 GAGAAAGAGCTGAATTAACAGAGATTTCTAAGCAAGAAATTTCTTTTGGAGTCGAG 900
QY 6524 TCGCAGAGCCGCGGTAGAGAGCCGAGAGTGGCAGAAACGTTTCGTAATCAATTTGAG 6583
DB 901 TCGCAGAGCCGCGGTAGAGAGCCGAGAGTGGCAGAAACGTTTCGTAATCAATTTGAG 960
QY 6584 GAAAGACGAGTAGATGTGGAAGACCAAGATTTTGTATAAAAAGTGAATTACG 6643
DB 961 GAAAGACGAGTAGATGTGGAAGACCAAGATTTTGTATAAAAAGTGAATTACG 1020
QY 6644 CAATAAGTTATAGTTCTGTGTGAATTTGAAATGCGCAGTGGAGTGGTGAAGAGAG 6703
DB 1021 CAATAAGTTATAGTTCTGTGTGAATTTGAAATGCGCAGTGGAGTGGTGAAGAGAG 1080
QY 6704 CTGGAATTGAGACAGAGTGGAGTTACCAAAAGATGAATCAACAGAGCAGAGTGA 6763
DB 1081 CTGGAATTGAGACAGAGTGGAGTTACCAAAAGATGAATCAACAGAGCAGAGTGA 1140
QY 6764 AAAATTTTAAATTTATGACTGCAAAACAGTTATGATGTATAGCAGAAAAATGA 6823
DB 1141 AAAATTTTAAATTTATGACTGCAAAACAGTTATGATGTATAGCAGAAAAATGA 1200
QY 6824 CAGGTAAGTGAATCGGTTACAGCCGAGCTGGAATTTCTTGACAGCCGAGATATCTGAGTGG 6883

```
Db 1201 CAGGACTGGAAATCGGTCAGCCGAGCTGGAATCTTGCAACCGAGATCTGAGTGG 1260
Qy 6884 TTTCTGTCAATTAATATGCAATTAAGSTGAACAAGATTCATCATGATCTTAACT 6943
Db 1261 TTTCTGTCAATTAATATGCAATTAAGSTGAACAAGATTCATCATGATCTTAACT 1320
Qy 6944 CTTCTACTGATGTAATGTAATGAAGCTCTTAATTAATTAATTCGAATTCCTTGAACGCGGTG 7003
Db 1321 CTTCTACTGATGTAATGTAATGAAGCTCTTAATTAATTAATTCGAATTCCTTGAACGCGGTG 1380
Qy 7004 GAGGAGCCGAGCTCTTGCAAGATTCACGAGTGGTTCTGTAACTATTAATAGTT 7063
Db 1381 GAGGAGCCGAGCTCTTGCAAGATTCACGAGTGGTTCTGTAACTATTAATAGTT 1440
Qy 7064 CTGTGATAGCTCGAGTTCACAACTGATTTGACCTCCGTAAGAGAAAAGTAATG 7123
Db 1441 CTGTGATAGCTCGAGTTCACAACTGATTTGACCTCCGTAAGAGAAAAGTAATG 1500
Qy 7124 TAACGCAAAAAGAGAAAAAATTAAGCAAAACAGCAATGCAAGATCGAGAG 7183
Db 1501 TAACGCAAAAAGAGAAAAAATTAAGCAAAACAGCAATGCAAGATCGAGAG 1560
Qy 7184 CAGCAATCGAGCCAAATGCTTGTGTAATTAATTTTGGAAACAGCTGTAGAGATAAAAA 7243
Db 1561 CAGCAATCGAGCCAAATGCTTGTGTAATTAATTTTGGAAACAGCTGTAGAGATAAAAA 1620
Qy 7244 ATTCTGAAGAGAAAGAAACAGAACTTTTAAACCTTTGACGAACTTAAACAAAGAAACAG 7303
Db 1621 ATTCTGAAGAGAAAGAAACAGAACTTTTAAACCTTTGACGAACTTAAACAAAGAAACAG 1680
Qy 7304 ATTAAGAAAGTAATGATGCTACGAAAAAATCTTCAATCAGCAGTATTTCTACAGAG 7363
Db 1681 ATTAAGAAAGTAATGATGCTACGAAAAAATCTTCAATCAGCAGTATTTCTACAGAG 1740
Qy 7364 ATACTTCTGTAAGAGCGATAGAGAGATCTCAGGAGAGAGAAATTAAGCCATTGGA 7423
Db 1741 ATACTTCTGTAAGAGCGATAGAGAGATCTCAGGAGAGAGAAATTAAGCCATTGGA 1800
Qy 7424 ACACCTTGATATTAATTTGAAAAAATGTAGATTAATTAACACAGAGACAAAGTAATATCA 7483
Db 1801 ACACCTTGATATTAATTTGAAAAAATGTAGATTAATTAACACAGAGACAAAGTAATATCA 1860
Qy 7484 CTTCTACTGATGTAATGTAATGAAGCTGGAACCTGCA 7510
Db 1861 CTTCTACTGATGTAATGTAATGAAGCTGGAACCTGCA 1887

RESULT 13
US-09-841-786-9
; Sequence 9, Application US/09841786
; Patent No. US2002005483A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO: 9
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-9
```

Query Match 11.6%; Score 1130; DB 3; Length 1130;

```
Best Local Similarity 100.0%; Pred. No. 8,9e-187;
Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCGCATCAAAAATTAACGTTACAGAGACAAAGAAAGAGATTCAGATTTCTAAAAA 60
Db 1 ATGAGCGCATCAAAAATTAACGTTACAGAGACAAAGAAAGAGATTCAGATTTCTAAAAA 60
Qy 61 GTTTTAATGATTTTGGAGATGTTGATTAACACTATGACGAGGCTATGATCAATC 120
Db 61 GTTTTAATGATTTTGGAGATGTTGATTAACACTATGACGAGGCTATGATCAATC 120
Qy 121 ACCGCACTGAGAAATTTTGGAAACAAAATAGAAAAAAGATTAATGTTATGACATTA 180
Db 121 ACCGCACTGAGAAATTTTGGAAACAAAATAGAAAAAAGATTAATGTTATGACATTA 180
Qy 181 ACCAAACAAGATTCAGAGGAGAAAGCTTTTAAACGTTTAAATAGATTTGCTTTAA 240
Db 181 ACCAAACAAGATTCAGAGGAGAAAGCTTTTAAACGTTTAAATAGATTTGCTTTAA 240
Qy 241 AATAATATAGCAAAATCTATATTTTGGGAAAAAGATAGTACGAGGCTAAATATCTTTT 300
Db 241 AATAATATAGCAAAATCTATATTTTGGGAAAAAGATAGTACGAGGCTAAATATCTTTT 300
Qy 301 AACTTTGTCAATGAAAAATTTGAAGTAGATGGAATTATCAACGAAATTCAGAAAAATAA 360
Db 301 AACTTTGTCAATGAAAAATTTGAAGTAGATGGAATTATCAACGAAATTCAGAAAAATAA 360
Qy 361 ATTGAGAGAAATTTATATTTTCTTAAGCTCGAGAGAGATGCGATAGAGAAAAATAGAG 420
Db 361 ATTGAGAGAAATTTATATTTTCTTAAGCTCGAGAGAGATGCGATAGAGAAAAATAGAG 420
Qy 421 ATCAATGCTGTTCTTTTCAATTTATTAATTCAGAAACAAAGATGATTTTAAAGAGCTT 480
Db 421 ATCAATGCTGTTCTTTTCAATTTATTAATTCAGAAACAAAGATGATTTTAAAGAGCTT 480
Qy 481 GAAGAGCCAAACATGTTAAAGTTTAAATGAAATCAATTCAGATAGATGAAAAAGTAAA 540
Db 481 GAAGAGCCAAACATGTTAAAGTTTAAATGAAATCAATTCAGATAGATGAAAAAGTAAA 540
Qy 541 ATTCATTTGAATTCGGAATGGAACATTCGATAGAGAAAAATCAATGCTGTTGAAGC 600
Db 541 ATTCATTTGAATTCGGAATGGAACATTCGATAGAGAAAAATCAATGCTGTTGAAGC 600
Qy 601 ATCGGTTTATATGCGCGGATATTAATGTAAGATTAAGTCAATTAATTAAGACAGAA 660
Db 601 ATCGGTTTATATGCGCGGATATTAATGTAAGATTAAGTCAATTAATTAAGACAGAA 660
Qy 661 ACAGATTTTAAAAATTTAGTCAATTAATTAAGTCAATTAATTTCTGTCGACCGAGAT 720
Db 661 ACAGATTTTAAAAATTTAGTCAATTAATTAAGTCAATTAATTTCTGTCGACCGAGAT 720
Qy 721 TTTAAAGCTACAGACAAAATCTGAGATATTAATCTTTAGCTCACATGATTTCTCT 780
Db 721 TTTAAAGCTACAGACAAAATCTGAGATATTAATCTTTAGCTCACATGATTTCTCT 780
Qy 781 CAAAAAGCTATGGGAAAAAATTCACATGTTGAAAAAGAAATAGAAATATATTAAGGA 840
Db 781 CAAAAAGCTATGGGAAAAAATTCACATGTTGAAAAAGAAATAGAAATATATTAAGGA 840
Qy 841 AATAACCAAGCAAAATTTGAATCTGATGCTGTAATTTGAAAGACATGAAAAATTA 900
Db 841 AATAACCAAGCAAAATTTGAATCTGATGCTGTAATTTGAAAGACATGAAAAATTA 900
Qy 901 AGTGGAAAGCTACAAATGGAAGATTTTAAGAAAGAGGAGAAAAAGAAATCTTAAC 960
Db 901 AGTGGAAAGCTACAAATGGAAGATTTTAAGAAAGAGGAGAAAAAGAAATCTTAAC 960
Qy 961 ACTCCTTAATGTTATACAGATGGAAGCTCGTAAGATTAATTAAGAAAGATCA 1020
Db 961 ACTCCTTAATGTTATACAGATGGAAGCTCGTAAGATTAATTAAGAAAGATCA 1020
Qy 1021 GAAAAGATGTTGACATTAACGCTGAAGCAAAAGATTTCTATGATGCAATTTAGTTACT 1080
Db 1021 GAAAAGATGTTGACATTAACGCTGAAGCAAAAGATTTCTATGATGCAATTTAGTTACT 1080
```

```
Db      1021 GGAAGAATGTGACATTACAGCTGAGCAAGAAATTTCTAATGATGCAACTTACTTACT 1080
Qy      1081 AAGCTTGCAAAAGCACTCTTTTACGTTTGTATGACAGGTTCTATTTCTCTAT 1130
        |||
Db      1081 AAGCTTGCAAAAGCACTCTTTTACGTTTGTATGACAGGTTCTATTTCTCTAT 1130

RESULT 14
US-10-647-057-9
; Sequence 9, Application US/10647057
; Publication No. US20040047871A1
; GENERAL INFORMATION:
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION
; APPLICANT: NAGARAJA, T.
; APPLICANT: STEWART, GEORGE
; APPLICANT: NARAYANAN, SANJEEV
; APPLICANT: CHENGARPA, M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: PCT-30962
; CURRENT APPLICATION NUMBER: US/10/647,057
; CURRENT FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-10-647-057-9

Query Match      11.6%; Score 1130; DB 7; Length 1130;
Best Local Similarity 100.0%; Pred. No. 8.9e-187;
Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  ATGAGCGGATCAAAATATACGTTGAGACAGCAAGAAAGGATATCAGATTCTAATAAAA 60
        |||
Db      1  ATGAGCGGATCAAAATATACGTTGAGACAGCAAGAAAGGATATCAGATTCTAATAAAA 60

Qy      61  GTTTTAATGATTTGGGATTTGTTGATTAACACTATGACGGTGAAGGCTTAATGATCAATC 120
        |||
Db      61  GTTTTAATGATTTGGGATTTGTTGATTAACACTATGACGGTGAAGGCTTAATGATCAATC 120

Qy      121  ACCGGACCTGAGAAATTTTGGAAACAAATAATGAAAAAAGATTAATGATTAATGATTAAT 180
        |||
Db      121  ACCGGACCTGAGAAATTTTGGAAACAAATAATGAAAAAAGATTAATGATTAATGATTAAT 180

Qy      181  ACAAACAAGATTCAGAGGGAGAAACGCTTTTACAGTTTAATGATTAATGATTAATGATTAAT 240
        |||
Db      181  ACAAACAAGATTCAGAGGGAGAAACGCTTTTACAGTTTAATGATTAATGATTAATGATTAAT 240

Qy      241  AATAATATAGCAAAATCTATATTTTGGGAAAAAGATATGATCGGGGTAAATATCTTTT 300
        |||
Db      241  AATAATATAGCAAAATCTATATTTTGGGAAAAAGATATGATCGGGGTAAATATCTTTT 300

Qy      301  AACTTTGTCATGAGAAAAATGAGTAGATGAGATTAATCAACGGAATTCGAGAAAAATATA 360
        |||
Db      301  AACTTTGTCATGAGAAAAATGAGTAGATGAGATTAATCAACGGAATTCGAGAAAAATATA 360

Qy      361  AATTGAGGAAATTTAATTTCTTAAGCTGGAAGGATAGGAGTGAAGAAAAATGAGATT 420
        |||
Db      361  AATTGAGGAAATTTAATTTCTTAAGCTGGAAGGATAGGAGTGAAGAAAAATGAGATT 420

Qy      421  ATCAATGCTGTTCTTTCTATCTATATTTCCAAAAACAAGATGATTTTAAAGAGGCTTTG 480
        |||
Db      421  ATCAATGCTGTTCTTTCTATCTATATTTCCAAAAACAAGATGATTTTAAAGAGGCTTTG 480

Qy      481  GAAAGAACCAACATGCTAAAGTTTAAATGAGATATTCAGTAGATGAGAAAAAGTAAAA 540
        |||
Db      481  GAAAGAACCAACATGCTAAAGTTTAAATGAGATATTCAGTAGATGAGAAAAAGTAAAA 540

Qy      541  ATTCCATTAATCGAATGGAAGCATTAAGGTGAGAAAGAAAAATCAATGCTGTTGAAGG 600
        |||
Db      541  ATTCCATTAATCGAATGGAAGCATTAAGGTGAGAAAGAAAAATCAATGCTGTTGAAGG 600
```

```
Qy      601  ATCGGTTATATGCGGCGGATATATAGATTGAAAGATCTGCAATCTAAAGACAGAAAT 660
        |||
Db      601  ATCGGTTATATGCGGCGGATATATAGATTGAAAGATCTGCAATCTAAAGACAGAAAT 660

Qy      661  ACAGATTTTAAATTTTATAGTCAATATTAATGATGATCAATTAATTTCTGCTCGACCGAGAT 720
        |||
Db      661  ACAGATTTTAAATTTTATAGTCAATATTAATGATGATCAATTAATTTCTGCTCGACCGAGAT 720

Qy      721  TTTAAAGCTTACCAAGACAAATCTGAGATATTAATTTCTTACGCTCAGATGATTTCTCT 780
        |||
Db      721  TTTAAAGCTTACCAAGACAAATCTGAGATATTAATTTCTTACGCTCAGATGATTTCTCT 780

Qy      781  CAAAAGCTATGCGGAAAAATTTCAACTGTGAAAGAAATGAAATGATTAAGAAAG 840
        |||
Db      781  CAAAAGCTATGCGGAAAAATTTCAACTGTGAAAGAAATGAAATGATTAAGAAAG 840

Qy      841  AATACCAAGCAAAATATTAATCTGATCTGATTAATGAAAGCAGATGAAATTAATAAAT 900
        |||
Db      841  AATACCAAGCAAAATATTAATCTGATCTGATTAATGAAAGCAGATGAAATTAATAAAT 900

Qy      901  AGTGGAAAGCTACAAATGGAAGATTTAATAAGAAAGGGAAGAAAGAACTTATAAC 960
        |||
Db      901  AGTGGAAAGCTACAAATGGAAGATTTAATAAGAAAGGGAAGAAAGAACTTATAAC 960

Qy      961  ACTCTTTAATGTTATATGATGATGAGATCTCCGTAAGATTAATTAAGAAAGATCA 1020
        |||
Db      961  ACTCTTTAATGTTATATGATGATGAGATCTCCGTAAGATTAATTAAGAAAGATCA 1020

Qy      1021  GGAAGAATGTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTACTTACT 1080
        |||
Db      1021  GGAAGAATGTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTACTTACT 1080

Qy      1081  AAGCTTGCAAAAGCACTCTTTTACGTTTGTATGACAGGTTCTATTTCTCTAT 1130
        |||
Db      1081  AAGCTTGCAAAAGCACTCTTTTACGTTTGTATGACAGGTTCTATTTCTCTAT 1130

RESULT 15
US-09-742-096-2
; Sequence 2, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DRUIHE, PIERRE
; APPLICANT: DAUBERESIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 5361
; TYPE: DNA
; ORGANISM: P. falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(5361)
; OTHER INFORMATION:
US-09-742-096-2

Query Match      1.2%; Score 116.2; DB 3; Length 5361;
Best Local Similarity 41.7%; Pred. No. 9e-10;
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

Qy      5126  GTTCGATGAACGGAAGCTTAGTTAGTATTCGAGTTGAAGAGTAAAGTTCTTCA 5185
        |||
```

Db 773 GTGTGAAGAAAGTATAGCTTCAAGTGTGATGAAAGTATAGATTCAAGATTGGAAGAA 832
Qy 5186 ATGTAGATGCAAAAGATCAAAAACATTAATAACA---TTGCCGGAATATCAATGAG 5242
Db 833 ATGTAGCTCCAAAGTGTGAAGAAATCGTAGCTCCAAAGTGTGTGAAGATGTGGCTCCA 892
Qy 5243 GAAAACGGCTGAGATTGAGACACAGTTGCTCATACAAATATGGAAGAAACATCAGTTA 5302
Db 893 GTGTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGTAGCTGAAATATGTGAAGAA 952
Qy 5303 TAGCATATGTAAGAAACAGTAAATTTACACGGCGCAATGATCAAGATGAAAGAAATATCA 5362
Db 953 GTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAA 1012
Qy 5363 ATGTAGCTGCAAAAGATTATATCTATGACCAATACTATAGCAGTCGAGATTGAGAGCA 5422
Db 1013 ATGTGAAGAAATCGTACTCCACTGTGTGAAGAAATGTGTAGCTCCAACTGTGAAGAA 1072
Qy 5423 AAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTACCTGATTAAGACAGTTTCTTCTC 5482
Db 1073 TTGTAGCTCCAAAGTGTGTAGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTAGAA 1132
Qy 5483 ATGTGTATCAAAAGTATATGACAAAGATTGAGAGAAAGAAATATGAAATTAAGAA 5542
Db 1133 ATGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAA 1192
Qy 5543 AAGCAATGTTAATGTCTAGCTGAAATACGAGTCAAGTGTCAACAATGCGACAGTGC 5602
Db 1193 GTGTAGCTGAAATGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAA 1252
Qy 5603 TTTCCGAGCAAGTGTGACAGAGCTGTGAGAGCTGTGAGTACAGTTAATAAATTACAC 5662
Db 1253 ATGTGAAGAAATCGTACTCCAACTGTGTGAAGAAATCGTAGCTCCAACTGTGAGAA 1312
Qy 5663 AAAATACCTCTGCACATATAAAAAATAGTACTCAAAATGTACGAAATCTTGTGAA 5722
Db 1313 TTGTAGCTCCAAAGTGTGTGAAGAGTGTGGCTCCAAAGTGTGAAGAAAGTGTAGAA 1372
Qy 5723 GC---AAATCTCATTCATCTATTAACAATTGCAATGAGCTGTGAGCTGTGAGCTGAG 5779
Db 1373 ATGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAA 1432
Qy 5780 GAGCTGAGTACAGAGTTCTGTAGCAGTGAATAGAGTTGTAATATATCGATAGCAGAT 5839
Db 1433 GTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAA 1492
Qy 5840 TAAATCATGCAAAATATCACTGCAAGGGAATGTTCGAGTTATTAACAGAGCTGATGGG 5899
Db 1493 ATGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAA 1552
Qy 5900 TAATTTGCTAATTTATGCAAGAAACAGTGTCTGAGAGTGGCCGTGACAGCAATTAGAGCTCA 5959
Db 1553 TCGTAGCTCCAAAGTGTGAAGAAATCGTAGCTCCAAAGTGTGAAGAAATTTGAGCTCCA 1612
Qy 5960 CCAGTGTGATGAAATTAACAGAGTCTACAAAAGCATATGTAAAGATTCTACAGTATG 6019
Db 1613 GTGTGTGTGAAGAGTGTGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTGAAGAA 1672
Qy 6020 CTAAGAGAAAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6079
Db 1673 GTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAA 1732
Qy 6080 AAGTATTCAAAATCTTAATTAATTAACGAAGACTTATCAAAAAGAAAGAAATTAAGTAATA 6139
Db 1733 ATGTGAAGAAAGTGTAGCT-----GAAATGTGTGAAGAAATGTGTAGCTCCAACTGTG 1786
Qy 6140 AAAAAGATTTGTATACCAATAGTCAAGTACTCATACTTTAAATCTTTATTTGCAATG 6199
Db 1787 AAGAAATGTAGCTCCAACTGTGTGAAGAAATTTGAGCTCCAAAGTGTGTA---GAAAGTG 1843
Qy 6200 CCGCTGTGTCAAGACACCGGAGAGTGTGAGAGACCTGTTAATATCAACAAGTTATGAG 6259
Db 1844 TGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAATG 1903

Qy 6260 AAACAGAGCTCTGTAGAAAATCTATATTAATGCAAAAATTAATCTGTAAATTCAG 6319
Db 1904 TTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAATCG 1963
Qy 6320 GAGATTACAGCAATTCAATCGAGTAGTAGGTTCTGTTGTGTGTGTGGAATGTAGAG 6379
Db 1964 TAGCTCCAACTGTGAAGAAATCGTAGCTCCAACTGTGTGAAGAAATGTAGCTCCAACTG 2023
Qy 6380 TAGAGCTCTCTGTATCCAAATTAATTAAG---AAATACCAAGCAAGTGTGAA 6436
Db 2024 TTGTGAAGAAAGTGTGCTCCAACTGTGTGAAGAAAGTGTGAAGAAATGTGAAGAAATG 2083
Qy 6437 AAATACCAATGTCTGATGAAGAGTTTCGAGGAAGAGCTGAATTTACAGCAATTTAGC 6496
Db 2084 TAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAATG 2143
Qy 6497 AAGGAATTTCTCTTTTGTAGTGTGAGTGTGAGAGCCGGGTGTAGAGCCGAGTGTGAG 6556
Db 2144 TTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAATGTAGCTCCAACTGTGTGAAGAAATG 2203
Qy 6557 GAACCGTTTCGTAAATCAATTTGACAGAAAGAGAGATGATGTGGAAGAGCAAG 6616
Db 2204 TAGCTCCAACTGTGAAGAAATGTAGCTCCAACTGTGTGAAGAAAGTGTGCTCCAACTG 2263
Qy 6617 TTTGTGAAGAAAGTGTGAGATTACAGCAAAAGCTTATAGTCTGTGCAATTTGGAATG 6676
Db 2264 TTGAAGAAAGTGTGAAGAAATGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAATG 2323
Qy 6677 CCGAGTGTGAGTGTGCTGCAAAAGAGCTGGAATTTGAGACAGCAGTGTGCAATG 6736
Db 2324 TAGCTGAAAGTGTGAAGAAAGTGTAGCTGAAAGTGTGAAGAAAGTGTAGCTCCAACTG 2383
Qy 6737 ATGAATCAACAGAGCAAGAGTGAAGAAATCTTAAGT---ATGACTGAAACAACT 6793
Db 2384 TTGAAGAAATTTGAGCTCCAACTGTGTGAAGAAAGTGTAGCTCCAACTGTGAAGAAATG 2443
Qy 6794 TAGATGTATGCAAGAAATAGATTAATTCAGATCTGGAATCGTTCAACCGAGCTG 6853
Db 2444 TTGCTGAAAGTGTGCAACAAATTTATCAGCAATCTTTAAGTAATTTATTAAGTGTGTA 2503
Qy 6854 GAATTTCTTGACCGGAGATCTGAGAGTGTCTGTCTGTCAATTAATTTGCAATTAAGTGTG 6913
Db 2504 TCGAACTGTAGGAATTAAGAGCAGTATATTAATGTAGATGAAGAGTAAAGAAATG 2563
Qy 6914 AAACAGATATGATCATAGTACTTTACACTCTTCTAGATGATAATGTAAGCTCTTA 6973
Db 2564 TAGTCAACCAATATCTAGAAAGTGTGAAGAACTACAGCTGAAAGTGTAACTTATTA 2623
Qy 6974 ATAAATTTTGAATTCCTTGACAGCCGAGTGTGAGAGCCGAGCTCTTGACAGATTACCG 7033
Db 2624 GTAACTATTTAGAGAGATTAACAAGAAATCTATTAATTAATTAATTAATTAATTAAG 2683
Qy 7034 GAGTGTCTTCTGTTAACATTAATAATGTTCTGTGATAGCTCGAGTTCAATTAATCTG 7093
Db 2684 TAGAAGAACTCCACGAAATGTATTAAGTGTGAGCTTTAAGAAATTAACCAAGTGAAGAG 2743
Qy 7094 ATTTGACTTCGTCTGCGAG---AAAGTGAATGTGAACGCGAAAGAGGAAATTAATTA 7150
Db 2744 AAAAGAAAGATTAATGATGTAATTAAGAAAGTGAAGAAAGTGTGCTTCACTTAA 2803
Qy 7151 AGCAAGACAGCAATGCAAGAAATCGAGAGCAGCAATTCGAGCAATGTCTTGTGTA 7210
Db 2804 TAGAAGATGTGTGAGACAGAGCAAGAAAGAGCGCAATTAACAATTAACGGAATATTTGAA 2863
Qy 7211 ATAAATTTGGAACAGCTGTGAGAGATGAAGAAATTTCTGAAGAAAGAAACAGAGTTT 7270
Db 2864 ATTTAGAAAGAAATGACAGTATGAAAGTATGAAATTTGTGACAGAAATTTAGAGAAATTA 2923
Qy 7271 TAAAACTTTAGACGAAGTTTAACAAGAACAGATTAAGAAATTAAGTAAAGTCTACGAAA 7330
Db 2924 ACGAACTGTATTTTAATCTGTATTAAGTAAGAGAGAAACGTAGAAATTAAGCGAG 2983


```

1 PRIOR APPLICATION NUMBER: PCT/FR96/00894
2 PRIOR FILING DATE: 1996-06-12
3 PRIOR APPLICATION NUMBER: FR 95/07007
4 PRIOR FILING DATE: 1995-06-13
5 NUMBER OF SEQ ID NOS: 29
6 SOFTWARE: PatentIn version 3.1
7 SEQ ID NO 1
8 LENGTH: 6152
9 TYPE: DNA
10 ORGANISM: P. falciparum
11 IS-05-742-096-1

```

Query Match	1.2%	Score 116.2	DB 3	Length 6152
Best Local Similarity	41.7%	Pred. No. 9.6e-10		
Matches 1747	Conservative	0	Mismatches 2373	Indels 67
				Gaps 14

QY	5126	GTTCTGATGAAACCGAGCTTTAGTTAGTATGCTTGAGTTTGGAAGAGTAAGTCTTTCA	5189
Db	1017	GTGTTGAAGAAAGTAAGTCTTCAAGTGTGATGAAAGTAATGATTTCAAGTATTGAAGAA	1076
QY	5186	ATGTAGATGCAAAAMAGATCAAAAACAAATAAATACAA---TTGCCGAAATGCAATGAG	5242
Db	1077	ATGTAGCTCAACTGTGTGAAGAAATCGTAGCTCAAAGTGTGTAGAAAGTGCGCTCCA	1130
QY	5243	GAAGAAGCGCTGAGATTGAGACAAAGTTGCTCATACAAATTTGGAAAAACATACATTA	5302
Db	1137	GTGTTGAAGAAAGGTAGAGAAATGTTGAAAGAAAGTGTAGCTAAAAATGTTGAAGAA	1196
QY	5303	TAGCTATTGTAAAAAACAAGTAAAAATTACACGCGCAATGATCAAGATAGAAAAATATCA	5366
Db	1197	GTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGTAA	1255
QY	5363	ATGTGACTGCAAAAGATTATATATGACCAATATCTATAGCAATCGAGATTGAGAGCA	5422
Db	1257	ATGTTGAAGAAATCCTAGCTCCAACTGTTGAAAGAAATCGTAGCTCCAACTGTTGAAGAA	1316
QY	5423	AAGAGAGCTCTGTGTGCAAGAGCTTGTGCAAGTACTACCTTGAAATAAGACAGTTTCTC	5487
Db	1317	TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA	1376
QY	5483	ATGTTGATCAAACTGATATATGCAAGAATTTAGAGAAAGAAATTAAGAAATAAGAA	5542
Db	1377	ATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA	1433
QY	5543	AGCCAAATGTTAAATTTCTAGCTGAAAAATGCAAGTCAAGTGTCACAATGTGACAGAGTC	5602
Db	1437	GTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAA	1496
QY	5603	TTTCCGAGACAAAGTGACAGAGCTGAGTAGAGAGCTGAGTATATAAATTACAC	5662
Db	1497	ATGTTGAAGAAATCGTAGCTCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAA	1555
QY	5663	AAAATACTTCTGCACATATAAAAAATAGTACTCAAAATGTACGAATGCTTTGTAAAA	5722
Db	1557	TTGTAGCTCCAAAGTGTGTAGAAAGTGTGCGCTCCAAAGTGTGAAGAAAGTGTAGAGAA	1616
QY	5723	GC---AAATTCATTCATCTATTAAACAATTTGAAATTTGAGCTGAGAGTGTGAGCTGAG	5778
Db	1617	ATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAA	1676
QY	5780	GAGCTGAGTGCACAGTTCTGTAGCAGTGAATAGATTGTTAAATATATCGATAGCAGAT	5833
Db	1677	GTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAA	1733
QY	5840	TAAATCATGCAAAAAATCACTGCGAAGGAAATGTGCAATTAATACAGAGTGTGATCGCG	5899
Db	1737	ATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAA	1796
QY	5900	TAAATGCTAAATATATGACAGAAACAGTGTGAGAGTGCCCGACAGATATAGAGCTCAA	5955
Db	1797	TGTTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCA	1855
QY	5960	CCAGTGTGAATGAATAATACAGAGCTACAAAGACATATGTAAAGATTTCTACAGTGAATG	6013

Db	1857	GTGTTGTGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAA	1916
Qy	6020	CTAAAGAAAGAAACAGATGATTATATTACTACAAAGGCAAGTAGATAAAGTGTAGATA	6079
Db	1917	GTTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA	1976
Qy	6080	AAGTATTCAAAAATCTTAAATATTACAGAAAGCTTATCACAAAAAAGAAAAATTAAGTAATA	6133
Db	1977	ATGTTGAAGAAAGTGTAGCT-----GAAAAATGTTGAAGAAATGCTAGCTCCAACTGTTG	2030
Qy	6140	AAAAAGATTTGTTATACCAATAGTTCAGCTACTCATCTTTAAATCTTTATTGCAAAATG	6198
Db	2031	AAGAAATGTAGCTCCAACTGTTGAAGAAATTTAGTCTCCAAAGTGTGTA--GAAAGTG	2088
Qy	6200	CCGCTGGTTCAAGACAAAGCCGAGTGGCAGGAACTGTTAATATCAACAAGGTTATGAG	6255
Db	2088	TGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAATG	2147
Qy	6260	AAACGAAGCTCTGTAGAAAATCTATATTAAATGCAAAACATTATCTGTAATATAG	6318
Db	2148	TTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATCG	2207
Qy	6330	GAGATTACAGAAATTCAATCGGAGTAGAAGGTTCTGTTGAGTGGTGGTGAAGATGAGAG	6376
Db	2208	TAGCTTCAAGCTGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATTTAGTCCAAAGTG	2266
Qy	6380	TAGAGCTTCTTCTGATACCAATATTATAAAG--AAATACCAAGCAAGAGTTGGA	6436
Db	2268	TTGTGAAGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTG	2322
Qy	6437	AAACTACAATGTCGTATGAAGGTTTCGAGAGAAAGCTGAAATTCACAGATTCTAAC	6496
Db	2338	TAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATG	2387
Qy	6497	AAGGAATTTCTCTTTGGAGTCCGAGTCGAGCAAGCCGAGGTGGAACCGGAGTGGAG	6556
Db	2388	TTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATGTTAGTCTCCAACTGTTGAAGAAATCG	2447
Qy	6557	GAACCGTTCCGTAATCAATTTTCGAGAGAAAGACGAGTAGATGTGAAAGCAAGAA	6616
Db	2448	TAGCTCCAACTGTGAAGAAATTTAGTCTCCAAAGTGTGGAAGAAAGTGTGGCTCCAAAGTG	2507
Qy	6617	TTTTGTGTAAGAAAGCTGAGATTACAGCAAAACGTTATAGTTCGTGCAATTGGAATG	6676
Db	2508	TTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTG	2567
Qy	6677	CCGAGTGGAGTGGCTGCAAAAGGAGTGGAAATTTGAGAGCAGCTGGCAAGTTCCAAAG	6733
Db	2568	TAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAACTG	2622
Qy	6737	ATGAATCAAAACAGAGACAGAGTGAAGAAATTTCTTAAATT--ATGACTCGAAACAGT	6793
Db	2638	TTGAAGAAATTTGTAGCTCCAAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAATG	2688
Qy	6794	TAGATGTATAGCAGAAAAATGAGATAAATCAGGTACTGGAATCGGTTCAAGCCGAGCTG	6855
Db	2688	TTGCTGAAAAGCTGCAACAAATTTATCAGCAAACTTTTAAAGTAATTTATTAAGTGGTA	2747
Qy	6854	GAATTTCTTCAGCCGAGATATCTGGAGTGTTCGTCAATATATTTGCCAAATTAAGTAG	6913
Db	2748	TCGAACCTGAGAAATTAAGACAGTATATTAATGAGATGAAGAAAGTGAAGAAATG	2807
Qy	6914	AAACGATATGATCATATGATCTTTACACTCTTACTAGATGTAAATGTAAGAGTCTTA	6972
Db	2808	TAGTCACCAATATAGAAAAACGTGAAGAAACCTACAGCTGAAAAGTGTAACTACTTTTA	2866
Qy	6974	ATAAATTTTCCAAATTCCTTGACAGCCGCTGAGAGCCGAGGTCCTTGACAGAGTTACCG	7033
Db	2868	GTAACATATTAGAGAGATACAGAAATTAATCTATTACTAATGATATCTATAGAGGAAAT	2922
Qy	7034	GAGTGGTTTCTGTTAACATATTAATAGTTCTGTGATAGCTCGAGTTCACAATTAAGCTG	7093

Db 2928 TAGAAGATCTCCAGCAAAATGATTAAGTCCCGCTTTAGAAAATACCAAGGTGAAGG 2987
Qy 7094 ATTTGACTTCGTACGAGA---AAAAGTAATGTAAACGGCAAAAGAGAAAAATATTA 7150
Db 2988 AAAAGAAAAGTAATAGTGAATTTGAAAGAGTAAAAAGAGGTGCTACCACTTTAA 3047
Qy 7151 AGCAACAGCGCAAAATCGAGGAATCGAGGACGCAATCGAGCCAAATGCTTGGTAA 7210
Db 3048 TAGAAGACTGTGAACAGGAGAAAGAAAGCGCAAAATCAATTACGGAATATTTGAAA 3107
Qy 7211 ATAAATTTGGAACAGCTGTGAAGATAGAAAAAATTTGAAAGAAAAGCAAGATT 7270
Db 3108 ATTTAGAAAAGAAATCGAGTGAAGATGAAGAAATGTTGCAAGAAATTTAGAGAAATTA 3167
Qy 7271 TAAAACTTTAGACGAAGTTAACAAGAACAGATAAAAAGTAAATGATGCTACGAAA 7330
Db 3168 ACGAACTGTATTTAACTGTATTAAGATAAGAGAAAAGATAGAAATTAACGCGAG 3227
Qy 7331 AAATCTTACATCAGCAGTATTTCTACGAAAGATCTCTGTAAGCGGATAGAGAG 7390
Db 3228 AAAGTTTGAAGAAACAATGAATGATTAAGCATTTTAAAGTAAATTTGATTAATGTA 3287
Qy 7391 ATACTCAGGAGAGAGAAATTAAGCATGTGAAGACTCTGATATTTAGAAAAAATG 7450
Db 3288 AAGGAATACAGAAAAATTTATTAACGATAG-----TTGGAAGTATGAACCACTA 3341
Qy 7451 TAGATTTTACAAACAGAGACAAGAAATATATCACTTCACTGCTGTTGGAACTGCAG 7510
Db 3342 TAGATATCAATCAGAGAAAGGTTGATTTGAAAGAAATGCGTTAGTTCGATTTTGA 3401
Qy 7511 GTCTGCTTCCGATCAGAACAGTGGCGCTTACAAATTTAAAGAAATTCGAGATTA 7570
Db 3402 ATATAATAGAAAAATGAAAGAGGTTTATTAATTAATTAAGAAATTTCAAGATACG 3461
Qy 7571 CTGTGAAAAATCTTTTGTGAAGCAGCTGAAAAGTAATGTTAGATGCGATAT----- 7625
Db 3462 AAGGTGTTCAAGAACTGTAACTGAACATGTAGAACAAATATGTAATGTGAGATGTTGAG 3521
Qy 7625 ---TACAGAAATGTTGCTTTAAACAGCATATCAAGCTCTGTAGAGCAATTTGGAAATG 7681
Db 3522 TTCCGTCTAGAAAGATCAATTTTAGGAATATTAATGAGGAGAGGTTGAAGAAA 3581
Qy 7682 GAGCTGCTATGCAAAATTAATTTCTAAATGGAAGATCAATATCAGATATTAATAATTTCTA 7741
Db 3582 TGTATTTTATTTGGAAGATGTAATTTAAAGTGAAGTGAATGTAATTAATCTGTAGAGAAA 3641
Qy 7742 AGTATTAAGAAAAAATTTGATGTAATGTAATGAAGATAATCGAATTTGAGCGGAG 7801
Db 3642 TTAAAGATGAACCGTTCAAAAAGAGGTGAAGAAAGAACTGTTAGTATTTGAAGAAA 3701
Qy 7802 CAAGAAGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTAATCTCAAAAGCAAAAGATG 7861
Db 3702 TGAAGAAAAATATTTGATGATGTAATTAAGAGAAAAAGAAAGTTTAAACGACAAATGA 3761
Qy 7862 AAATGAATTCAGAGTTGAAATTTGAGAAGATATTTTCAATGAAGAAAATAGATTAATCTA 7921
Db 3762 TAGATCAGTAGAAGATCAATGAAATATCTTCAATTTCAAAAGAAATCGAATTTCTA 3821
Qy 7922 GCCCTTCTAAAGAAATTTGAGAAGAAATCAATCTCAAGTGAAGAAAAGAAACAGATGA 7981
Db 3822 TTAAAGATTAAGAAAAAGATGTTTCACTAGTTGTGAAGAGTTCAAGCAATGATATG 3881
Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGCAGTACGAGGCGAG 8027
Db 3882 ATGAAAGTGTGGAAGAGTTTGAATGGAATATGGAAGAGAGATTAATGAAGATG 3941
Qy 8028 AATTAATTTCCGAGCAAAAGATCCGAGACTCTTAATTTGAAGTTAGTACAAATCCG 8087
Db 3942 CTGTTAAATTAATTAATGACTTACTAGCAAACTTAATTTGAAGAACTCAAGAGTTTAAAGAG 4001
Qy 8088 AAGAAATTTTCAAGCAGATATATGGAATATGGAAGCAACATTAATAAGAAATGAAAGTAA 8147
Db 4002 TAGAAGCAGATTAATTAAGATGATGAAAAATTTAAAGATTAAGAAAAAGC--ATTATC 4059

Qy 8148 AGCAATTTCTAAGCAGTAAACAGGTTCTGTATTGAGAGAGATGGAGTCAACCAAGCAGA 8207
Db 4060 AGAAGATCTAAGAAATTAATGATGCAAAAGATATCAATTAAGAAAAATTTGAAGA 4119
Qy 8208 AGTACTGCTCAGGTAAACATAAGTGAAGTGAAGAGAAATTTGTTCAAGACAA 8267
Db 4120 GGAACATGATATTAACGAGCGTTGATGAAGTTGATTAATTAAGATGTCGAAGAGA 4179
Qy 8268 TCGATTTGATCAATTTTAAAGTGAAGGTTTGAATGAAGTAAAGTAACTGCTAAATC 8327
Db 4180 CAAGATCGAAAAAGATCTGATTTAAAGATCTTGAAAGATATTAATTAAG----- 4232
Qy 8328 TTCTGTAGTATCAGAAATGAGAGAAATGCGGACGAGAGATGATTAATCTTACAGC 8387
Db 4233 --AGTAAAGAAATCAAGAACTTTGAAGTGAATTTTGAAGATTTAAAGATTAA 4290
Qy 8388 ACAAGTAAATCTGATCCGTAGTTCTTTACGAAGCAAGTTATGAAGATTAATGATTA 8447
Db 4291 AACTTTGAACAGATATTTTGAAGAGAAAAAGAAATTAAGAAAGATCAATTTGAAA 4350
Qy 8448 CACAAAAAATATATTTCAAGATCAATGCTTGTCTTTAAATGATACAAAGATGAAGC 8507
Db 4351 ATTCGAAGAGAGCTGAAGAAATTAAGATCTTGAAACAGATATATTTAAAGAGTATC 4410
Qy 8508 GAATATGAATCTTACGCGTACCGCGTGCATCAAGAGAACAAACAAAGCATTTAC 8567
Db 4411 TTCAATTAAGTTGAAGAGAAAAAATTAAGAAAGTACAGATTTAAAGAAAGAGT 4470
Qy 8568 GAGATCAACAAAGTTAACCTTCACTCACTGAATGAGAAACGTATCTCACTTCGTC 8627
Db 4471 AGAACATATTAATAGTGTGATGCCATATTAAGGTTTGAAGAAAGATGATTTGAAGA 4530
Qy 4531 AGTATAGTGA-TTTAAAGAGAGTATATTAAGATGTTAAAGGAGATATGAG--ATTAG 4587
Db 8628 AAAAGCTTGGCTAAAAATGAAATTAATGAAGTAAAGTAAAGAACTGAGAGCCTTACT 8687
Qy 8688 CGAGCGGAAACAGACCGCTTGAATTAATCAAGAGTACTACAGAGCTTGGTTC 8747
Db 4588 GGATATGATTAAGAAAGTTTGAAGATGTAAACAAACTTGAAGAAAGGTTGAATC 4647
Qy 8748 AGAAATTTGGAAATTTGAGATTAATTAAGAAACGATTCGAAAGATTAATCAAGATTGTAAG 8807
Db 4648 CTTAAAGATGTTTATCTAGTCATTTAGGCTATGATGAAGAACAAATGAAGAAACAGAAA 4707
Qy 8808 ACTCAACGAGACGAAACCAAGAGGCTTGTGAAAGAAATGTAATTTCTGTGAAAA 8867
Db 4708 AAAAGCTCAAGACCTAAGTTGGAAGATTAATTAAGAAAGAGTTAAAGAAAGACC 4767
Qy 8868 TACAATTTGAGGGAACAAATCACTCAATGAAGTAAAGCAGAAATTTGCAACCG 8927
Db 4768 AAAAGAAAAAATTAACAAAAAGAAAGTAAAGTTGATTAATTAAGTAAAGAAACAAAGAA 4827
Qy 8928 AAGTGAATGTAAGTCTTGAATGAATGATGATGATCTTCAAGAAAAAGTGTG 8987
Db 4828 TCAATATGTAAGGTTGAATGAATGAATGAATTAAGAAAGATGTAAGAAAGATAT 4887
Qy 4888 AGAAGAGATTAAGAAAGATTAAGTTGAATTAAGTGAAGATTAAGTGAAGATAT 4947
Db 9048 AGCCAAAATGAGAGACATGTAATGTAAGAACTACTGCAAAACAAAGATTAACAGATT 9107
Qy 4948 AGGTGAAGCAAGAGTGAAGTTATGATTAATGCTCAAAAGAGAAACGCAATGAAAA 5007
Db 9108 TACAAGAGCAAGTAATTAATCTTGAAGAGAGCGCTGACGCTGACGTAATATC 9167
Qy 5008 GGTAAAGGAAAAAGAAAAATTAAGAAAAAGTTGAAGAGGTGTTAGTGTCTTAA 5067
Db 9168 GATGTACATTTCCATGAGATGATTAATAATTTGGCAAGCAGTATGATCTTC 9227
Qy 5068 AAAACAGTGAAGAGATATGAATATATGTTCAAAAAATTTGATTAAGAAAGTTGATTAAGA 5127

```
Qy 9228 TCAATTAAACCAAAATTCAAAATAATATTAATTACTTAGCATCAA 9274
Db 5128 AGTATCTAAAGCTTTAGAAATCAAAAATGATGTTACTAATGTTTAA 5174

RESULT 17
US-10-415-253-1
; Sequence 1, Application US/10415253
; Publication No. US20040067236A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Joe
; APPLICANT: Drulline, Pierre
; TITLE OF INVENTION: Immunogenic Compositions Comprising
; TITLE OF INVENTION: Liver Stage Malarial Antigens
; FILE REFERENCE: B45250
; CURRENT APPLICATION NUMBER: US/10/415,253
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: PCT/EP01/12349
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: EP00203724.0
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5528
; TYPE: DNA
; ORGANISM: Kl Parasite Strain
US-10-415-253-1

Query Match 1.1%; Score 107.6; DB 7; Length 5528;
Best Local Similarity 41.6%; Pred. No. 2.9e-08;
Matches 1741; Conservative 0; Mismatches 2379; Indels 66; Gaps 14;

Qy 5126 GTTCGTGAAGAAACGAGCTTGTAGTGAATTCGAGTTTGAAGAGTAAGTTCTTCA 5185
Db 941 GTTTGAAGAAAGTATAGCTTCAAGTGTGATGAAGAACTATGATTCAGATATTGAAGAAA 1000
Qy 5186 ATGTGATGCAAAAAGATCAAAAACAAATTAATACAA---TTGCCGAAATGCAAAATGAG 5242
Db 1001 ATGTAGCTCCAACTGTGAAAGAAATCGTAGCTCCAAAGTGTGTGAAGAAATGCGCTCAA 1060
Qy 5243 GAAAACCGCTGAGAGTTGAGCAACAGTTGCTCATACAAATATGGAAGAAACATCAAGTTA 5302
Db 1061 GTGTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA 1120
Qy 5303 TAGCTATTGTAAAAACAGTAATAATTACACCGCAATGATCAAGATGAAAAATATCA 5362
Db 1121 GTGTAGCTGAATAATGTTGAAGAAAGTGTGAAGAAATGTTGAAGAAATGTTAGCTGAAA 1180
Qy 5363 ATGTGACTGCAAAAAGATTACTATGACCAATACTATAGCAGTCCGAGTTGAGAGCA 5422
Db 1181 ATGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAA 1240
Qy 5423 AAGAGAGCTCTGTGCAAGAGCTTGTGCAAGTACTACTGTAATGAAGAAGTTCTTCTC 5482
Db 1241 TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAA 1300
Qy 5483 ATGTGATCAAACTGATATTGCAAGATTTAGAGAGAAAGAAATATAGAAATTAAGAAA 5542
Db 1301 ATGTGGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA 1360
Qy 5543 AGGCAAAATGTATGTTCTAGCTGAAAATATAGAGTCAAGTCAAGT--GCAGACAGT 5600
Db 1361 GTGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1420
Qy 5601 GCTTTCGAGAGCAAGTGAAGAGCTGAGTAGAGAGCTGAGTAGCAAGTTAATTAATTAAC 5660
Db 1421 TGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAAT 1480
Qy 5661 ACAAATATCTTGCACATATAAAAAATAGTACTCAAAATGTACGAAGATGCTTGTGTAA 5720
Db 1481 TGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGTGAAGAAAGTGTGAAGAAA 1540
```

```
Qy 5721 AAGCAATTCATTCATCTATTAAACAATTGGAATTGAGCTGAGAGTTGAGCTGAGG 5780
Db 1541 TGTGGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAG 1600
Qy 5781 AGCTGAGTGCAGAGTTCTGTAGCAGTGAATTAAGATTGTAATTAATACGATAGCAAAATT 5840
Db 1601 TGTAGCTGAATAATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1660
Qy 5841 AAATCATGCAAAAATTCATGCGAAGAGCAAAATGTGCGAGTTATTAACAGAGTGTAGTCCGT 5900
Db 1661 TGTGGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAT 1720
Qy 5901 AATTGCTAATTATGCAAGAACAGTGTGTGAGTGCGCCGTCAGCAATAGAGCTCAAC 5960
Db 1721 CGTAGCTCAACGTGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAAAG 1780
Qy 5961 CAGTGTGAATGAATTAACAGATCTCAAAAGCATGTGTAAAGATTCTACAGTGAATTGC 6020
Db 1781 TGTGTGAAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTAGAGAAATGTTGAAGAAAG 1840
Qy 6021 TAAAGAAAGAACAGATGATTATATTACTACTCAAGCGCAAGTAGATAAGTGTAGATTA 6080
Db 1841 TGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1900
Qy 6081 AGTATTCAAAATCTTATATTAATAAGAAAGTGTATCACAAAAAGAAAAATAGTATTA 6140
Db 1901 TGTGGAAGAAAGTGTAGCT-----GAAAATGTTGAAGAAATCGTAGCTCCAACTGTTGA 1954
Qy 6141 AAAAGATTTGTGTACCAATAGTTCAGCTCATACCTTAAATCTTTATTTAGGCAATATC 6200
Db 1955 AGAATCGTACTCCAACTGTGAAGAAATGTAGCTCCAAAGTGTGA---GAAAGTGT 2011
Qy 6201 CGCTGCTTCAAGAACAGCGGAGTGTGCGAGAACTGTAAATATCAACAAAGTTTATGAGA 6260
Db 2012 GCTTCCAAGTGTGGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGT 2071
Qy 6261 AACCAAGCTCTGTGAGAAAATCTATATTAATGCAAAAACATTATCTGTAAATCAGG 6320
Db 2072 TGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATGT 2131
Qy 6321 AGATTACGCAATTCATCGAGATGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6380
Db 2132 AGCTCCAACTGTGAAGAAATCGTAGCTCCAACTGTGMAAAATGTTAGCTCCAAAGTGT 2191
Qy 6381 AGAGCTTCTCTGATACCAATATTAATAAG---AAATACCAAGACAGAGTTGAGAA 6437
Db 2192 TGTGAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGT 2251
Qy 6438 AACTACATGTGTGATGAAGGTTTCGAGAAAGCTGAAAATTAACAGCAGATTCTAACCA 6497
Db 2252 AGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGT 2311
Qy 6498 AGGAAATTTCTCTTTTGAAGTGTGAGTGTGCGAGTGTGCGAGGCTAGAGCCGAGTGTGCG 6557
Db 2312 TGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGT 2371
Qy 6558 AACCGTTTCGTAATCAATTTGACAGAAAGACGGAAGTAGATGTGAAGAAAGAAAT 6617
Db 2372 AGCTCCAACTGTGAAGAAATGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGT 2431
Qy 6618 TTTGTAAAAAAGCTGAGATTACAGCAAAAAGTATAGTTCTGTGTGCAATTTGAAGATGC 6677
Db 2432 TGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGT 2491
Qy 6678 CGCAGTCCGAGTGTGCAAAAAGAGCTGGAATTTGAGACAGCAGTGTGCAATTAACAAAGA 6737
Db 2492 AGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAACTGT 2551
Qy 6738 TGAATCAAAACAGAGAGAGTGAAGAAATTTCTAAAT---ATGACTCGAAACAAGTT 6794
Db 2552 TGAAGAAATTTAGCTCCAAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGTGT 2611
Qy 6795 AGATGTATATACAGAAATAGATTAATAATCAGATCTGAAATCGGTTCAAGCCGAGAGCTG 6854
```


APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 4761
LENGTH: 7434
TYPE: DNA
ORGANISM: Staphylococcus aureus
US-09-815-242-4761

Query Match 0.8%; Score 79.2; DB 3; Length 7434;
Best Local Similarity 41.3%; Pred. No. 0.0029;
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

QY 3782 TTCAGATGAAACAAGTGTCTTGTTAAATATGATGATTCGAAATGCCATTAATTTA 3841
DB 3659 TTTATATTAAGAAAGCACTGCACTGATGAAATCAAGCAATTTTAAATACAAATTCAG 3718
QY 3842 AAGTAATGCTTTGAGTGAGGAATCAAGTACAGACAGACAGAGCTTTGAAACAGTTA 3901
DB 3719 AATTCAGACTACGCCAGATGCAACAGATGAAGAAAAACAGCAGCTGATGCGAAGCAA 3778
QY 3902 AAGAAAGTGAGACAGAAAGAAAGTATCTATTTGGAACTTCTGTTCTATCACTTGG 3961
DB 3779 AATCTGAAATGATTAAGCAATTCAGCACTTTCAGCAGCACTATTAACGCAAGTTG 3838
QY 3962 TGAACATGATGTTCTGCAAAATCAGAAATATATCACTAGACAGAGATCTGAAAGCC 4021
DB 3839 ATGAACTTAAAGCAATGCAAGACGCAATTAATCGGTACACCAAAAGTTGTGAAGA 3898
QY 4022 AAAAAATGATGTTGATGTCACTGCTTATCAAGCGGACACCCAAAGTGAAGAGCTTTAA 4081
DB 3899 AACAGCGGCTAAGATGAATGATCAATTCAGCAACGCAACAAATGTTATCAATA 3958
QY 4082 ATTTACAACTGGAAGTCAATGGAATCTAGAGGCTACTGTGATCTTTGCCAATTTAA 4141
DB 3959 ATGATCAGAACGCTACACAGAAAGAAAGAACACACTTATTCACAAATTAGCAACGCG 4018
QY 4142 ACAACAAGTAATGCTTCTATTAAGTGTGAGGATATATCTTAACGTTAATCGAGCGGACG 4201
DB 4019 TTAACAGACCGAAAAATTAATTAACGTGCACTGATGATTAATGTTGATCAGGCGCA 4078
QY 4202 CAAGAAGCTTTTATGCAACCACTCAAGTACTGCTCAGTGAAGCAAGGAGGAGCAATTA 4261
DB 4079 AAGAGCTGGAAGAAATTCATTTCAAGGACCGCAACAGCAACAGGCTTTAATCAATG 4138
QY 4262 GTTCTGAGCGGATTAAGAAATTAATCAAGGCGCTGTTCTGTCAATTAAGATTGAAGG 4321
DB 4139 CTAAATATGATGTTGATCAAGCTGTGCAACTCAAAATCAAGCAATGTAATCAACTG 4198
QY 4322 AGGTGAAGCTAGCTGATTAATCTTCCATGGAAGAGCTAATGAATCATGTCAATG 4381
DB 4199 GTGCTACAACTGAAGAAAAATGACGCAAAAGATTTAGTTTAAAGCTTAAGAAAAAG 4258

QY 4382 CCAAGATGCAAGAAAGTCTGATCTAGCAAAAGAAATATCAGGCTTACTAATATGAA 4441
DB 4259 CGTATCAAGATATCTTAAATGCAAAACACTTAATGATGTTACGCAATTTAAAGTCAAG 4318
QY 4442 AAGATAAAAATATTTAGAAAGATCGTGTATTTAATACACTGGAATGTTATTAACGA 4501
DB 4319 CAGTTGCGATATTCAGAGTATTACTGCAATACAAATTAAGATGTTG-----CGA 4372
QY 4502 AGGAACAATGAAAAAGCAAAAGAAAAAGAGCGGCTCATTTAATGCTGCTTAT 4561
DB 4373 AAGATGAATTTAGCAAAAGCAAAAGCAAAAGAGCGCTTATTTGCAAAATGCAAGATG 4432
QY 4562 CGGTTGCGAAGGATTAATCCGCTGAGAGGATGATGCTATTCAGTCAATACTGTTAAA 4621
DB 4433 CGACTACTGAAGAAAAAGAACAGCAAAATCAACAGTGAAGCAGCAATTAACACAAGTA 4492
QY 4622 ATAAATTTAAGCAAAATTTGAGTGGAACCAATTAAGAAAGCCGAGAGGATTAATTCATG 4681
DB 4493 ATCAAAATATTAAGAAATGCAAGTCAATCGATGATGAACACTGCAAAAGATTAATGCA 4552
QY 4682 CGAAACATGTAATGAGGAGGCAAAATCATCTACTGTTGTTGTAATGCGCTTGTGAC 4741
DB 4553 TTCAAGCAATTCACCAATTCAGAGCATCAAGATGTTAAAAGAAATGCAAGAGC-GGA 4611
QY 4742 TTGCTATGAGAAAGATGCTTTTCAAGAAATGGGATCTGAGCATGCGCAAGCTTATCA 4801
DB 4612 TTGCTACTGAAGTCAAAATTAATTAATCTGAATATCTTAATTAATGAGCT----- 4665
QY 4802 ATGACACATTCGAAGGTGATTAAGAAAGAAATTTCTGCTGATTCCTTAAATGTAAGC 4861
DB 4666 -----ACTAATGAAGAAAGAAAGTAAAGATTTGACAGCTTGAACAGATTAAGAAAG 4720
QY 4862 CAATTAATTCATCTTGGGAGTAATGTTGCGGAAACCAATGCGGTTCTCTTTCTACGG 4921
DB 4721 GTTAAATTAATTAATGATGACAGCACTACTACAGTATGTAATCTACTGTAAGATACG 4780
QY 4922 CGGTGAGAGCTCTTTGCGAATATCTCTTCAATTAATTAATCTGCTTTGATTAACG 4981
DB 4781 CAGTCAAAAGATTCAACAACTTCATGCAAACTCTTTAAGAAACAGAGGTTAAAG 4840
QY 4982 GAACAGGTAATCTTTTATGTAAGAAATATCAAAAGTCAATGATCAAGCTTTGAATG 5041
DB 4841 AATTAATCAAGCTCAGCTGATTAAGAAACAAATTAAGAAACCAAAATGATCAAC 4900
QY 5042 ATTTCAATTAACAAAGCTTTCTGCTGAGGCGCTGCAATTAAGCAAGCTGGAATG 5101
DB 4901 AACAAAGATTAATGATGCAAAACAAAGATGATGATTAATCAAGCGAAACAA 4960
QY 5102 GAGGAATGATCTGTCAATCGGTCTGATGAACGGAACCTTAATGATGATTCG 5161
DB 4961 ATGTGATCAATCAATCAAAATGAATGTTGATTAATGCAATTAAGAAAGAAAGCTA 5020
QY 5162 AGTTGAAGGATTAAGTTCTTTCAATGATGATCAAAAGATCAAAACAAATTAATCA 5221
DB 5021 AATTAATGACGTTTAAACATTTAGTGAATCAAAAGAAAGTCTTTAGTAAATTTGAAG 5080
QY 5222 TTGCGGAAATCAATGAGGAAAGCGGCTGAGGTTGAGCAACAGTTGCTCATCA 5281
DB 5081 ATGCAATTAATGCTAAAGTAAAGGAGGATTAATCTTAACCATGCACTTCAAGTGA 5140
QY 5282 ATATTGAAACAAATCAATTAATGCTATTTGTAAGAAACGTAATTAACAGCGCAATG 5341
DB 5141 TTGCTGAAGCGAAACAAACCTTGCTGAATTAACAACTCGGATCAAAATGTTAAATC 5200
QY 5342 ATCAAGATTAAGAAATATCAATGATGCTGCAAAAGATTAATCTATGACCAATTAATG 5401
DB 5201 AAGCTACTCTTAAGATGCAATTTGAAGTTCAATTAATCAATTAATTAATTAAG 5260
QY 5402 CAGTGAAGTTGAGAGGAAAGAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTACT 5461
DB 5261 ATTAACCAATTCACAGGTAAGAAAGATCAAGTACAGATTAATTAATGCTTAATGAG 5320
QY 5462 TGAATTAAGAGTTTCTTCTCATGTTGATCAACATGATTAATGAACAAAGATTAAGAGAG 5521

Db 5321 ATCAGAGAAAATATATATTTCAGCTGACACTAATGCAACAGATGAAAAAGCAACAG 5380
Qy 5522 AAAATATAGAAATAAGAAAAGCAAAATGTTAATGTTCTAGCTGAAAATAGAGTCAAG 5581
Db 5381 CAATTAAGCAAGTTGACAAAATG--TTCAAACTGCAATTAGAAAGCATTAATAATGGTG 5437
Qy 5582 TGGTCACAAAATGCGACAGTCTTTCGAGCAAGTGCAGCAAGCTCAGTAGAGCTGAG 5641
Db 5438 TGGATTAATGGTACGTTGATGATGCACTTAACCAAGTAAAGCAGAAATGATGCTATTC 5497
Qy 5642 TAGCAGTTAATAAATTACAAAATACTTTCGACATTAATAAAATAGTACTCAAAATG 5701
Db 5498 AAGTAGATGCTACTGTTAAACCTAAAGCGAACCAAGCTATTGAAGTTAAAGCAGAAATG 5557
Qy 5702 TAGCAATGCTTTGGTAAAG-----CAAATGCTAGTATCA 5740
Db 5558 CGAAAGATCTATTGATCAAAAGTGACCACTAATGCTGAGAAAAAACTGAAGCTTAG 5617
Qy 5741 TTAAACAATTGGAATTGAGCTGAGTGGAGCTGAGAGCTGAGAGCTGAGAGCTTCTG 5800
Db 5618 CAATGATTAAACAATTACAGATCAAGTAAACAAGTATTACTGATGCAACAACAGT 5677
Qy 5801 TAGCAGTGAATAAGTTGTAATATAGATAGCAAGATTAATATCATGCAAAATCATCTG 5860
Db 5678 CTGAAGTTGAAAAAGGAAAGCTCAAGACCTGAAGCAATTGATTAACATTCAAATGCACT 5737
Qy 5861 CGAAGGAAATGCGAGTTATTAACAGAGTCTGATGCGGTAATTGCTAATTAGCAGGA 5920
Db 5738 CAACGAAAAAACAAGGCTATCGAAGATTAGAAATCTGACTAGACAGATTGAAGCAG 5797
Qy 5921 CAGTCTGAGAGTGGCCGTGACGAAATAGAGCCTCAACAGTGAATGAATTAACAG 5980
Db 5799 GGTAAATGTCAACGCTGATGCTACAACTGAAGAAAAAAGAGCTTACAGATGCTTAG 5857
Qy 5981 GATCTCAAAAAGCATATGTAAGATTTCTACAGTATGCTAAGAGAAACAGATGATT 6040
Db 5858 AAGACATTTTCAAAAAGCACTGAAGATATTTCTGATCAAACTCAATGACAGAAATCG 5917
Qy 6041 AATATTACTCTCAAGGCGCAAGTAAAGTGTGATGAATAAGTATTCAAAAATCTTAATA 6100
Db 5918 CTACTGTCAAAAATAGTGGCTTGAACCACTTAAGCACACGATTAATCTGAAAGTTA 5977
Qy 6101 TTAAACAAGCTTATCACAAAAAGAAAAATAGTAAATAAAAAAGATTTGTTACCAATA 6160
Db 5978 AAAAAAATGCTTTGGAAGCAATCAGAGAGTGTTAACAAGCAATTAATAATTAATAA 6037
Qy 6161 GTTCAGTACTCACTT 6178
Db 6038 ATGCAGATGCAGATGCAT 6055

RESULT 20
US-10-470-048B-201
; Sequence 201, Application US/10470048B
; Publication No. US2005003744A1
; GENERAL INFORMATION:
; APPLICANT: MEINKE ET AL.
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
; FILE REFERENCE: SONN:035US
; CURRENT APPLICATION NUMBER: US/10/470,048B
; NUMBER OF SEQ ID NOS: 603
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 201
; LENGTH: 7434
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-470-048B-201

Query Match 0.8%; Score 79.2; DB 8; Length 7434;
Best Local Similarity 41.3%; Pred. No. 0.0029;

Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;
Qy 3782 TTCAAGTAAACAAGTCTTTGTTAAATAATAGATATTCGAATATCCATAAATTTA 3841
Db 3659 TTAATTAATAAAGCACTGCAAGTATTAATTAACAAGCAATTTTAATAACAAATTTGCAAG 3718
Qy 3842 AAGTAAATGCTTTAGTGAAGAACTCAAGTACGACGAGCAAGTTTGAAGCATTTA 3901
Db 3719 AGATTCAAGCTACGCGAGATGCAACAGATGAAGAAAAAAGAGCTGATCTGAAAGCA 3778
Qy 3902 AAGAAATGAGAGCAAGAAAGAAAGTTATCTATTGGGAACCTTCTGCTTCAACTTAG 3961
Db 3779 ATACTGAAATGTAAGCAATCAAGCATTTAGCAGCAACTTACACCAACAGTTG 3838
Qy 3962 TGAACATGAGTTTCTGCAAAATACGAATAATATACAGTACGAGAGAAATCTGAAGCC 4021
Db 3839 ATGAAGCTAAAGCAATGCAAGAGCAGGATTAATGCGGTAAACACCAAAAGTTGGAAG 3898
Qy 4022 AAAAAATGATGTTGATGTCAGTCTTATCAAGCGGACACCAAGTGAACAGAGCTTTAA 4081
Db 3899 AACAAGCGGCTAAGATGAATGATCAATTACAAAGCAACGCAACAAATGTTATCAATA 3958
Qy 4082 ATTTAACAAGCTGAAGATCAATGGAATGTAAGGGGCTACTGTGATCTGTTCCAAATTA 4141
Db 3959 ATGATCAAGAGCTTCAACAGAGAAAAAGAGAGCTATTCAACATTTACACACAGAG 4018
Qy 4142 ACAACAAATGAAATGCTTCTATTAGTGGTGGAGATTAATACTTAATGAGCGGAGC 4201
Db 4019 TTACAGAGCGGAAAAATTAATTTACAGCTGCACTGATGATTAATGTTATGAGGGA 4078
Qy 4202 CAAGAAGCTTTTATGCAACCACTCAAGTCTGACGTGACGAGCGGAGGACATTA 4261
Db 4079 AAGAGCTGGAAGAAATTCATTCATCAAGCAGCAACGCAACAGCGGTTAATCAATG 4138
Qy 4262 GTTCTGAGCGGAGTTAGAAATTTTCAAGGGCTGTTCTGTCAATTAAGTTGCAATG 4321
Db 4139 CTAAAAATGATGTTGATCACTGACACTGACAACTCAAAATCAAGCAATTAATATCACTG 4198
Qy 4322 ACGTGAAGCTAGGCTGATTAATCTTCATCGAAGAGCTAATGAATCAATGCTATTG 4381
Db 4199 GTGCTACACTGAAAGAAAAATGCAAGAAAGATTTAGTTTAAGCTTAAGAAAAAG 4258
Qy 4382 CCAAGATGTCAAAAGAAAGTTCTGATCTAGCAAAAAGAAATCAAGCTTTACTAAATGAA 4441
Db 4259 CGTATCAAGATATCTTAATGCAACAACTAATGATGTTAGCGAAATTAAGATCAAG 4318
Qy 4442 AAGATTAATAATTTAGAAAGTGTGATTTAATCAAGCTGGAATGTTATTAACA 4501
Db 4319 CAGTGTGATATTCAGATTAATCTGCAAGATTAACAATTAAGATGTTG-----CGA 4372
Qy 4502 AAGAACTAAGTAAAGCAAGCAAGAAAAAGAAAGAGCGGCTATTGTAATGCTGCTTAT 4561
Db 4373 AAGATGAATTAGCAACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4432
Qy 4562 CGGTGCTGGAACGATTAATCCGCTGAGAGGATGATTCAGTCAATCTGTTAAA 4621
Db 4433 CGACTACTGAAGAAAAAGAAAGCAAGCAATCAAGAGTGAAGCAATTAACAAGGTA 4492
Qy 4622 AATAATTAAGAGCAATTTAGAGTGAAGCAATTAAGAAACCGGAGAGATTAATTCATG 4681
Db 4493 ATCAAAATATTGAATAATCAACAGTCAATCGATGATTAACATGCAAAAGTAAATGCA 4552
Qy 4682 CGAAACATGTAATGAGGCAAAATCATCTACTGTTGTTGTAATGCGGCTTGGAGC 4741
Db 4553 TTCAAGCAATTTGACCAATTCAGAGCATCAAGATGTTAAACGAATCGAAGAGC-GGAA 4611
Qy 4742 TTGCTATCAGCAAGAGTCTTTTTCAGGAATGGAATCTGAGCATGCGCAAGCTTATCA 4801
Db 4612 TTGCTAATGAAATGCAAAATTAATAATTAATTAATTAATTAATTAATTAATTAAT 4665
Qy 4802 ATGACAGATTTGCAAGAGTGAATTAAGAGAAATTTCTGCTGATTCCTTAATGTAAGC 4861
Db 4666 -----ACTAATGAAGAAAAAGTAAAGATATTGACCAAGTTAGAGCAGCATATGAAGAG 4720

Db 3719 AGATTCAAGCTACGCGACAGATGCAACAGATGAAAGAAAAACAAGCAGCTGATGCTGAAGCAA 3778
 Qy 3902 AAGAAAGTGAGAGACAAGAAAAAGTTATCTATTGGGAACTTCGCTCTATCAACTAG 3961
 Db 3779 ATACTGAAAAATGTTAAAGCAAAATCAAGCCATTTCAGCGCAACTATCAACGACAAGTTG 3838
 Qy 3962 TGAACAAATGAAGTTCTGCAAAATGCAAAATATATACAGTAGCAGAGATCTGAAGCC 4021
 Db 3839 ATGAAGCTTAAGCAAAATGCAAGAGCGAATTATGCGGTAAACCAAAAGTTGTAAGA 3898
 Qy 4022 AAAAAATGATGTGATGCTACCTGCTTATCAAGCGGACACCCAGTAGACAGAGCTTTAA 4081
 Db 3899 AACAAAGCGCTAAACATGAAATTTGATTCATTTCAAGCAACGCAAAACAAATGTTATCAATA 3958
 Qy 4082 ATTCAAGCTGGAAGTCAAAATGGAACGTAGGGGCTGTTGCTGATCTGTGCAATTAA 4141
 Db 3959 ATGATCAAGACCTTCAACACAGAAAGAAAAAGACAGCTATTCAACAAATTAAGCAACGAG 4018
 Qy 4142 ACAACAAAGTAATGCTCTTATTAAGTGGAGAGATATATCAAGCTTAATCGAGCGGAG 4201
 Db 4019 TTACAGACGCGCAAAATATATATTAACGCTGCAACTGATGATATATGTTAGATCAGGCGA 4078
 Qy 4202 CAAAAGCTCTTTAGCAACCACTCAAGTACTGCTGCAAGTGAAGACGAGGAGGACAATTA 4261
 Db 4079 AAGAGCTGGAAGAAATTCATTTCAAGACGCAACGCAACAGCGGTTAATCAAAATG 4138
 Qy 4262 GTTCTGAGACGGGATTAGGAATTAATCAAGGGGCTGTTCTGCTAATAAGATTGACAAATG 4321
 Db 4139 CTAATAATGATTTGATTCAGAGCTGAGACCTCAAAATTCACAAATTAATATACACTG 4198
 Qy 4322 ACGTGAAGCTAGCGCTTGAATATCTTCATCGAAGAGACTAATGAATCAATGTCATTTG 4381
 Db 4199 GGTCTACACACTGAAGAAAAATGACAGCAAAAGATTAGTTTAAAGTAAAGAAAAAG 4258
 Qy 4382 CCAAAAGATGTCAAAGAACTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGAA 4441
 Db 4259 CGTATCAAGATATCTTAATATGCAACAACAACTAAATGATTTACGAAATTAAGATCAAG 4318
 Qy 4442 AAGATTAATAAATATTAGAAAGATCGATGATTAATATACGCTGGAATGGTATTATAGA 4501
 Db 4319 CAGTTGCTGATATTCAAGGTATTACTGAGATTAACAATTAAGATGTTG-----CGA 4372
 Qy 4502 AAGAACAACTAGAAAAAGCAAAAGAAAAAGAGAGCGGCTCATTTAAATGCTGCTTAT 4561
 Db 4373 AAGATGAATTTGCAACAAAGCAAGCAAAAGAAAGCGCTTATTGCACAACTGCGAATG 4432
 Qy 4562 CGGTTGCTGGAACGGAATTAATCCGCTGAGAGAGTACTATTGCACTCAATACTGTTAAA 4621
 Db 4433 CGACTACTGGAAGAAAAAGCAAGCAAAATCAACAAGTAGACGCAATTAACACAAAGTA 4492
 Qy 4622 ATTAATTTAAAGCAGAAATTGAGTGAAGCAATTAAGAAAGCCGGAAGAGATTAATAATTCATG 4681
 Db 4493 ATCAAAATATTGAATGCAACGTCAAATGATGATGAATTAACACTCAAAAGTAATGCAA 4552
 Qy 4682 CGAAAATGTAATGTGAGCGCAAAATCATCTACTGTTGTTGTGATGCGGCTTCTGAC 4741
 Db 4553 TTCAAGCAATTGACCCAAATTCAGAGATCAACAGATGTTAAACGAATGCAAGAGC-GGAA 4611
 Qy 4742 TTGCTATACGCAAGATGCTTTTTCAGAGATGGGATCTGAGCATGGAAGCTTATCAA 4801
 Db 4612 TTGCTACTGGAATCCAAAATTAATAATTAATCTGAATAATGAATGAGCT----- 4665
 Qy 4802 ATGAACAGATTGCAAAAGGTGATTAAGAGAAATTTGCTGATTCCTTAATGTGAAGC 4861
 Db 4666 -----ACTAATGAAGAAAAAGGTAACGATTTGGAACAGTTAGACACATATGAAGAG 4720
 Qy 4862 CAATAATTTCAATCTTGTGGGTGAATGTGCGGAAACATTTGCGGTTCTTTTACAG 4921
 Db 4721 GTTAAATATATATTAATGACGAACTACTACAGGTGATGATATCTGCTAAAGATACAG 4780
 Qy 4922 CGGTGAGAGCTGCTTTTGCAATTAATCTCTTCAATATAAACCTTCTGTTGATTACAG 4981

Db 4781 CAGTACAAAAAGTTCAACAACTTAGCAAAATCTGTTTAAAGAAACACAGAGTAAAAAAAG 4840
 Qy 4982 GAACGAAGGTAAATCCTTTTAGTGAAGAAATTAACAAAAGTCAATGTATCAAGCTTGAAATG 5041
 Db 4841 AATTAGATCAAGCTGACAGCTATTAAGAAAAACAATTAAGAACAAACACCAATGATCATAC 4900
 Qy 5042 ATTCATATTTACAAAGCTTTCGCTGAGAGCGCTGCAAGTATTAAAGAGCTGGAATCG 5101
 Db 4901 AACCAAGAAATTAATGATGCAAAAACAGAAAGTTGATATCTGAATTAATCAAGCAAAAAACA 4960
 Qy 5102 GAGGAATGATATCTGTCATCTGCTGTTGATGAACGGAAGCTTATGATGATTTG 5161
 Db 4961 ATGTGATCAATCAATCAACAAATGAATATGTTGATTAATGAGCTTAAGAAAGAAAGCTTA 5020
 Qy 5162 AGTTGAAGAGTAAAGTTCTTCAATGATGATGCAAAAGATCAAAAAACATTAATTAACA 5221
 Db 5021 AAATTAATGACGTTAAACATTTAGTGTGTCAAAAAGAAAGATGCTTAAATTAATGAAG 5080
 Qy 5222 TTGCGGAAATGCAAAATGAGAAAAAGCGGCTGAGATTGAGCAACAGTTGCTCATACA 5281
 Db 5081 ATGCAATTAATGCTTAAGTAAACGAAGCGATTAATCTTAAGCATCGACTTCAAGTGA 5140
 Qy 5282 ATATTGAAAAAATCACTTATAGCTATTGTAATAAACAAGTAAATTAACAACGCGAAATG 5341
 Db 5141 TTGCTGAAGCGAAACAAAACTTGCTGATTTAAACAAACTGCGGATCAAAATGTTAATC 5200
 Qy 5342 ATCAAGATAGAAAAAATATCAATGATGACTGCAAAAGTATATATAGCAATTAATCTATAG 5401
 Db 5201 AAGCTACTCTTAAGATGACATTGAAGTTCAAAATTCATTAATGACTTATATTAATTAACG 5260
 Qy 5402 CAGTGGAGTTGAGAGCAAAAGAGCGCTGTCGCAAGAGAGCTTTCGCAACTACTACT 5461
 Db 5261 ATTAACAATTTCAACAGGTAAAGAAAGATCAGCTACAAAGATTTATATGCTTATAGCAG 5320
 Qy 5462 TGAATTAAGACGTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAAGAGAG 5521
 Db 5321 ATCAAGAAAAAATTAATATTCACCTGACACTAATGCAACCAAGATGAAGAAAGCAACAG 5380
 Qy 5522 AAAATTAATGAATAATGAAGAAAAAGCAAAATGTTAATGTTCTAGTGAATAATCGAGTCAAG 5581
 Db 5381 CAATTAAGCAAGTTGACCAAAATG---TTCAAATCTGCTTGAAGAGCTTAATTAATGATG 5437
 Qy 5582 TGGTACAAATGCGACAGTGTCTTCCGAGCAAGTGAACAAGTGCAGTAGAGCTGAG 5641
 Db 5438 TGAATTAATGATGAGTTGATGATGATTAACAAGAAAGTAAGCAAGCATTTATGATATTC 5497
 Qy 5642 TAGCATTAATTAATAATTAACAAATACTTCTGCAATATTAATAAATATGTACTCAAAATG 5701
 Db 5498 AAGTGAATGCTACTGTTTAAACCTTAAGCAACCAAGCTATTGAAGTTAAAGCAGAAAGTA 5557
 Qy 5702 TAGAAATGCTTGTGTAAGAG-----CAAAATCTGATTCATCTA 5740
 Db 5558 CGAAAAGATCTATTGATCAAAAGTGAACAAGTTAATCTGTAAGAAAAAACTGAAGATTTG 5617
 Qy 5741 TTAACAAATTTGAAATTTGAGAGCTGAGAGTGAAGCTGAGAGAGCTGAGATGACAGTTTCTG 5800
 Db 5618 CAATGATTTAAACAATTAACAGATCAAGCTTAACAAGATTAATCTGATGACCAACAACACTG 5677
 Qy 5801 TAGCATGAATTAAGATTGTAATAATTAATGACATGCAAGAAATTAATCATGCAAAAAATCACTG 5860
 Db 5678 CTGAAGTTGAAAAAGCGAAAGCTCAAGAGCTTGAAGCATTTGAATTAACATTCGAATTCGACT 5737
 Qy 5861 CGAAGGGAAATGTGGAGTTTATTAAGAGCTGAGTGGGGTAAATGCTTAATTAATGACAGAA 5920
 Db 5738 CAACAGAAAAACAAGAGCTATCGAAGATTAAGAAATCGCACTGACCAAGATTGAAGAG 5797
 Qy 5921 CAGTGTGAGAGTGGCCGTCGACCAATAGAGAGCTCAACGAGTGTGATGAATGAATTAACAG 5980
 Db 5798 GTTAAATGATGAAGCTGATGCTTAACAAGTGAAGAAAAAGAGCGTTTACGAATGCTTTAG 5857
 Qy 5981 GATCTACAAAAGCATATGTAATAAGATTTACAGATGATGCTTAAGAAAGAAACAGATGATTT 6040
 Db 5858 AAGACATTTTATCAAAAGCAACTGAAGATATTTCTGATCAAACTACAAATGCAAGAAATCG 5917

QY 6041 ATATTACTACTCAGGGCAAGTAGATTAAGTGTGATTAAGTATTCAAAAATTCTTAATA 6100
DB 5918 CTACTGTCAAAAATAGTGGCTTGGAACAACTTAAGCCACMACGTATTAATCTGAAGTTA 5977
QY 6101 TTAACGAGACTTATACAAAATAAGAAAATAAGTAATTAAGAAATTGTACCAATA 6160
DB 5978 AGAAAAAGCTTTGAGACGATCAGAGAGAGTGTTAACAGCAATAGAAATTAATTAATA 6037
QY 6161 GTTCAGCTACTCATCTT 6178
DB 6038 ATCGAGTCGACATGCTAT 6055

RESULT 22

US-08-781-986A-63
Sequence 63, Application US/08781986A
Publication No. US20030054436A1
GENERAL INFORMATION:
APPLICANT: Charles Kunach
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 8155 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-63

Query Match 0.8%; Score 79.2; DB 2; Length 8155;
Best Local Similarity 41.3%; Pred. No. 0.003;
Matches 998; Conservative 0; Mismatches 1178; Indels 42; Gaps 5;

QY 3782 TTCAAAGTAAACAGTCTTTGGTTAAATAAGTATTTGCAATGCAATTAATTTA 3841
DB 5651 TTAATTAATAAAGCACTGACGTATGAAATTCACAGCAATTTTAATACAAATTGCAAG 5710
QY 3842 AAGTAATGCTTTGAGTGAAGAACTCAAGTACAGAGAGAGAGAGTTGGAAGCACTTA 3901
DB 5711 AGATTCAAGCTACCGCAGATGCAACAGATGAAGAAAAACAAGAGCTGATGCTGAAGCA 5770
QY 3902 AAGAAAGTGAGACAAGAAAAAGTATCTATTGGGAATCTTGCTTTATCAACTTAG 3961
DB 5771 ATACTGAAAATGTAAGGCAAAATCAAGCCATTTTCAGAGCAACTACTAAGCAACAAGTG 5830
QY 3962 TGAACAATGAAGTTTCTGCAAAATCAGAAAAATTAATACAGTAGCAGAGAAATCTGAAGCC 4021

DB 5831 ATGAAGCTAAAGCAAAATGCAAGAGAGGATTTAATGCGGTAAACACCAAAAGTTGGAAGA 5890
QY 4022 AAAAAATGATGTTGATGTCCTGTTATCAAGCGACACCCAGTGCAGAGCTTTAA 4081
DB 5891 AAACAAGCGGTAAAGATGAAATGATCAATTAACAAGCAAGCCAAACAAATGTTATCAATA 5950
QY 4082 ATTTACAGCTGGAAGTCAAAATGCACTGTAAGGGGCTACTGTGACTGTGCGCAATTTA 4141
DB 5951 ATGATCAAGCGTACCAACAGAGAAAAAGAGAGCTATTTCACAAATTAGCAACGAG 6010
QY 4142 ACAACAAGTAAATGCTTTCTAATTAAGTGGGAGATATACTAACCTTAATCAGCGGACG 4201
DB 6011 TTACAGAGCGCAAAAATTAATTTACAGCTGCACATGATGATATGTTGATGATGAGCGCA 6070
QY 4202 CAAAAGCTTTTTCACCAACTCAAGTACTGCTGACGTGACGACGCGGAGGACAAATTA 4261
DB 6071 AAGACGCTGAAAGAAATTCAAATTCAAACAGCAACAGCAACACCGGTTAATCAAAATG 6130
QY 4262 GTTCGAGCGGATTAAGAAATTTCAAGGGGCTGTTCTGTCAATAAGATTGACAAATG 4321
DB 6131 CTAATAATGATGTTGATCAAGCTGTGACAACTCAAAATCAACAAATGATTAATCAACTG 6190
QY 4322 AGTGAAGCTAGCGCTGTGATTAATCTTCATGGAAGAGCTAATGAATCAATGTCATYG 4381
DB 6191 GTGCTACCACTGAAGAGAAAAATGCAAGCAAAAGATTTAGTTTAAAGCTTAAGAAAAAG 6250
QY 4382 CCAAAAGATGTCAAGAAAGTTCTGATCTGCAAAAGAAATATCAGCTTTATCTAAATGCA 4441
DB 6251 CGTATCAAGATATCTTAATATGCAAAACAACTAATGATGTTACGCAAAATTAAGATCAAG 6310
QY 4442 AAGATAAAAATATTGAAGATCGTGTATTAATCACTGGAATGTTTATTAACA 4501
DB 6311 CAGTGTGATTAATCAAGTATTAAGTATGCAAGATCAACAAATTAAGATGTTG-----CGA 6364
QY 4502 ACGAACTAGAAAAAGCAAAAGAAAAAGAGAGCGGTCAATTAATGCTGCTTTAT 4561
DB 6365 AAGATGAATTTACCAACAAAGCAAAAGCAAAAGCAAAAGCGTTATTTGCAACAATGCGAGAG 6424
QY 4562 CGGTGCTGGAACGAGTAATTCGCTGAGAGAGTACTGATTCAGCTCAATCTGTTAAA 4621
DB 6425 CGACTACTGAAGAAAAAGAAACAAGCAAAATCAAGTAGACACACATTAACACAAGTA 6484
QY 4622 ATTAATTTAAACAGAAATTTGAATGAGCAATTAAGAAACCGAGAGGATTAATTCATG 4681
DB 6485 ATCAAAATATTGAATGCAAGTCAATGATGATGTAACACTGCAAAAGATTAATGCA 6544
QY 4682 CGAAACATGTAATGAGAGGCAAAATCATCTACTGTTGGAATGCGGCTTGTGAC 4741
DB 6545 TTCAAGCAATTTGACCAATTTCAAGCATCAACAGATGTTAAAAAGAAATGCAAGC-GGA 6603
QY 4742 TTGCTATCAGAAAGATGCTTTTTCAGAAATGAGATGAGATGAGACATGCAAGACTTATCA 4801
DB 6604 TTGCTATCAGAAATGAGAAATTAATAATTAAGTAAATTAATTAAGACT----- 6657
QY 4802 ATGACAGATTGCAAAAGTGAATTAAGAAAGATTTCTGCTGATTTCTTAATGTAACG 4861
DB 6658 -----ACTAATGAAGAAAAAGTAAAGATTTGACCAATTTGAGAGAGATGAAGAAAG 6712
QY 4862 CAAATTAATTCATTTCTGGGGTGAATGTTGCGGGAACCAATGCGGCTTCTTTTACAGG 4921
DB 6713 GTTTAATAATTAATTAATGAGCACTACTACAGTATGATTAATCTGTTAAAGATACAG 6772
QY 4922 CGGTAGAGCTGCTTTTTCGAAATTAATATCTTATTAATAAACCCTGCTTTGATTACAG 4981
DB 6773 CAGTCAAAAAGTTCACAAACTTCATGCAATCTGTTAAAGAAACAGCAGGTAAAGAAAG 6832
QY 4982 GAAAGAGTAAATCTTTTATGGAAGAAATTAACAAAGTCAATGATGACAGCTTTGAATG 5041
DB 6833 AATTAGATCAAGCTGAGCTGATTAAGAAAAACAAATTAAGCAACCAACAAATGATCAC 6892
QY 5042 ATTCTATTTTCAAAAGTTTCTGCTGAGGCGCTCAAGTATTAAGCAGGCTGATGATG 5101

```

Db      6893 AACAGAAATTAATGATGCAAAAACAGAAAGTTGATCTGAAATTAATCAAGCAAAAACAA 6952
Oy      5102 GAGGAATGATCTGTCAATCGTGGTCTGATGAAACGGAAGCTTTAGTATGATTCG 5161
Db      6953 ATGTGATCAATCATCAACAAATGAAATGTTGATTAATGCACTTAAGAGAAAAGCTA 7012
Oy      5162 AGTTGAAAGAGTAAGTCTTTCAATGTAGATGCAAAAAGATCAAAAACAAATTAATACAA 5221
Db      7013 AAATTAATGCAAGTTAAACATTTAGTATGATCAAAAAGATCTTTAGCTAAATTTGAAG 7072
Oy      5222 TTGCGGAAATGCAATGAGAAAAGCGCGCTGAGAGTTGAGCAACAGTGTCTATACA 5281
Db      7073 ATGCAATTAATGCTAAAGTAAACGAAAGGAAATCACTTAACGCACTGACTTAAGTGA 7132
Oy      5282 ATATTGAAAAACATCAAGTTATAGCTATTTGTAATAAATTAATTAACAGCGCAATG 5341
Db      7133 TTGCTGAAGCGAAACAAAACCTTGTGATTTAAACAACTCGGATCAAAATGTTAAATC 7192
Oy      5342 ATCAAGATAGAAAAAATATCAATGTGACTGCAAAAAGATTAATCTATGACCAATACTATAG 5401
Db      7193 AAGCTACTTCTAAAGATGACATTTGAAGTTCAAAATCATTAATGACTTAATATTAACG 7252
Oy      5402 CAGTGGAGTTGAGAGAGAAAGAGAGCTCTGTGCAAGAGAGCTTCTGCAATGACTACT 5461
Db      7253 ATTACACAAATTCACAGGTAAAGAAATGAGTACACAGATTTATATGCTTATGCAAG 7312
Oy      5462 TGAATTAAGCAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTATAGAGAA 5521
Db      7313 ATCAGAAAGAAAATTAATTTAGCTGACATTAATGCAACAGAAATGAAAAGCAACAG 7372
Oy      5522 AAAATTAATGAAAATTAAGAAAAGCAAAATGTTATGTTCTAGCTGAAAATACAGAGTCA 5581
Db      7373 CAATTAAGCAAGTTGACCAAAATG---TTCAAACTGCAATTAAGAAATTAATTAATG 7429
Oy      5582 TGGTCAAAATGCGACAGTCTTCCGAGAGCAAGTGACAGCAAGTCAAGTGAAGTGGAG 5641
Db      7430 TGGATTAATGAGCGCTTATGATGATTAATTAACCAAGGTAAAGCAACAAATGTTGCTAT 7489
Oy      5642 TAGCAGTTAATTAATTAACCAAAATTAATCTTGCAATTAATAAATAGTACTCAAAATG 5701
Db      7490 AAGTGAATGCTACTGTTAACTTAAGGCAACCAAGCTATTGAAGTTAAAGCAGAAAGTA 7549
Oy      5702 TAGCAATGCTTTGTTAAAG-----CAATCTCATTCATCTA 5740
Db      7550 CGAAAGATCTATTGATCAAAAGTGAACAGTTAACTGCTGAAGAAAACGTAAGCATTAG 7609
Oy      5741 TTAACAATGATGATTTGAGAGTGAAGTGAAGCTGAGAGAGCTGAGATGACAGGTTCTG 5800
Db      7610 CAATGATTAACCAATTAACAGATCAAGCTAAACAGGTAATTAATGATGCAACACAGCTG 7669
Oy      5801 TAGCAGTAATTAAGATTGTAATTAATTAAGATTAAGATTAATTAATGCAAAATCACTG 5860
Db      7670 CTGAAGTTGAAAAGCAAGAGCTCAAGAGCTTGAAGCATTTGATTAACATTCCTCAATG 7729
Oy      5861 CGAAGGAAATGTCGAGTTATTAACAGATCTGATGCGGTAAATTCCTAATTAATGAGANA 5920
Db      7730 CAACAGAAAACAAAAGCTATCGAAAGATTAGAAATGCACTAGACCGAGATTGAAGCAG 7789
Oy      5921 CAGTGTCTGAGATGCGCCGTGACAGCAATGAGAGCTCAACAGGTGATGAATTAACAG 5980
Db      7790 GTGTAATGTCACCGCTGATGCTACAACTGAAGAAAAGAAAGACGTTTTCGAAATGCTTAG 7849
Oy      5981 GATCTACAAAAGCATATGTAAGAAAGATTCTACAGTATGCTTAAGAGAAAACAGATGAT 6040
Db      7850 AAGACATTTTATCAAAAGCACTGAAGATATTTCTGATCAAACTCAATGCAAGAAATCG 7909
Oy      6041 ATATTACTACTCAAGGCAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA 6100
Db      7910 CTACTGTCAAAAATAGTGGCTTGAACAACCTTAAGCAACAGATTTATATCTGAAAGTTA 7969
Oy      6101 TTAAGGAAGACTTATCACAAAAAGAAAATTAAGTATTAATAAAGAGATTTGTTACCAATA 6160
Db      7970 AGAAAAATGCTTTGGAAGCAATCAGAGAGTGTGTTAACAGCAAAATAGAAATTAATTA 8029

```

```

Oy      6161 GTTCAGTACTCATACTT 6178
Db      8030 ATGCAGATGCAGATGCAT 8047

RESULT 23
US-10-329-624-63
; Sequence 63, Application US/10329624
; Publication No. US20040043037A1
GENERAL INFORMATION:
APPLICANT: Charles Kunach
            Gil H. Choi
            Patrick S. Dillon
            Craig A. Rosen
            Steven C. Barash
            Michael R. Farnon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/329,624
FILING DATE: 27-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248PDL1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 8155 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-10-329-624-63

Query Match      0.8%; Score 79.2; DB 7; Length 8155;
Best Local Similarity 41.3%; Pred. No. 0.003;
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

```



```

Db      7970 AGAAAAATGCTTGGAGCAATCAGAGAAAGTGTAAACAGCAATAGCAATATTAAAA 8029
Qy      6161 GTTCAGTACTCATCTT 6178
Db      8030 ATGCAGATGCAGATGCAT 8047

RESULT 24
US-09-815-242-4580
; Sequence 4580, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 4580
; LENGTH: 7104
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4580

Query Match      0.8%; Score 74.8; DB 3; Length 7104;
Best Local Similarity 41.6%; Pred. No. 0.017;
Matches 992; Conservative 0; Mismatches 1367; Indels 27; Gaps 7;

```

```

Db      3002 AACGTAAACT---GCAATTTGAACCAATGAATGATTGCACTATGGAAGAACAAAGCTG 3058
Qy      7568 TTACTGTGAATAATCTTTTGTGAAGAGCTGAAAAAGTAATGTTAGATCGATATTA 7627
Db      3059 CAAAAGATTAAGTTGATCAAGACAGTATTACTGCAAGCGTGAATATGATATGCTGAG 3118
Qy      7628 CAGAAATGTTGCTTTAAACAGCATATCAAGGTCCTGTAGAGCATTTGGAAATGAGGCTG 7687
Db      3119 CAAATAGTATGATGATATTAATGCAAAAACTACTAATGAAAGTCAATCGACCATTTACAC 3178
Qy      7688 CCTATGCAATTTAATTTCTAATGAGAGATCAATATCATGATTTAAAAATTTCTAAGCTAT 7747
Db      3179 CTGATGCAAAATGTTAAACCAACAGCGAAAACAAGCAATTCGTAAATGATACAAAGCGCAAG 3228
Qy      7748 TAGAAAAAATATTGATGTTA---TTGTAAAAAGATTAATTCGAAATTTGAGACGCGAAGCAA 7804
Db      3239 AACGCAATGATGCTATTAATACGTTGCAACACAGAAAGAAAAAGCAGCTCGGAAACAC 3298
Qy      7805 AAGATTAAACCGTAGAGACGCTAGCTGCGAGCCATTATCTCAAAAGCAAAAGATGAA 7864
Db      3299 AAGTTCAAACTGAAAAAACAACAGCTGATACAGCAATTTGATGTGCAATACAAATGCG 3358
Qy      7865 TGAATTCAGAGCTTGAATTTGAAGAGTATTTTCATGAAAGAAATGAGTAACTAGCC 7924
Db      3359 AAGTTGAAGCGCTAAAAAATCAGAAATTCGTAATTTGAAGCAATTCAGCCAGCAACAA 3418
Qy      7925 CTCTTAAAGGAATG---GAAGAGAAATCAATGTCTAAAGTGGAAAAAGAAAAACAGAGTGA 7981
Db      3419 CACTTAAGATTAATGCGAAACAGCAATTCGTACGAAAGCGAAATGAAGTTAAACAGCAA 3478
Qy      7982 CTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGCAGGCGCAGAAATTTTCCGAG 8041
Db      3479 TCGCTCAAAACCGAAGACATTACTGCTGTAAGAAATTTGACGCGCAATGCGAAT---GTAG 3535
Qy      8042 CAAAAGATGCCGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAGTATTTTC 8101
Db      3536 ATAAATGCTGTACCAACAGCAATTAACAACATTTGAAGCTGTAATAGTCAAAATGATGAG 3595
Qy      8102 ATGCAGATTAATGTGAATTTGAGCAACACATTAATAATGAAGTAAACAGCAGTTCTTAAAG 8161
Db      3596 ACCAAGCAAAACACATGCTGTAAGTATGATTTGATCAAGTAAACACACAGTTATTAAGA 3655
Qy      8162 CAGTAAACAGTTCTGTATTTGAGAGAGTTGAGATCCACAGCAGAACTACTGCTGAG 8221
Db      3656 AAGCAACAGCAGTTAACAAGACGCAAAAATTAATTTACAGCTGCAACTGATATATAGTG 3715
Qy      8222 GTAAAACTATGTGTAAGTTGAGAGAAATTTGTTCAAGAACAAATGATGAAATGCA 8281
Db      3716 TAGATACAGCAAGAGATGCAAGTAAAAATTCGATTCAAAGTACACACACACAGCGG 3775
Qy      8282 TTTCTAAAGTAGAAGTTTGGATGAAAGTAAAGTAACTGCTAAATCTTCTGTAGATCAG 8341
Db      3776 TTAAATCAATGCTAAAAATGATGTTGATCAAGCTGTGACAACTCAAAATCAAGCAATG 3835
Qy      8342 GAAATGAGAGGAATTTGCCGAGCAGAGATGAAATCTTTCAACAGCAAAAGTATATCTG 8401
Db      3836 ATATATCAACTGCTGCTACACTGAAGAAAAAATGACGCAAAAGATTTAGTTTAAAG 3895
Qy      8402 AATCCGTAGTTCTTTACGAAGCAAGATTATGAATAATATGATTTACAAAAAATATA 8461
Db      3896 CTAAGAAAAAGCGTATCAAGATATCTTAAATGCAAAACAATCAACATGTTGCGCAA 3955
Qy      8462 TTTGAGAAGTGAATGCTCTTGCTTTAATGATTAAGAAATGAAGCGAATPAGAACTCT 8521
Db      3956 TTAAAGATCAAGCAGTTGCTATGTTCAAGGTTATTTCTGCAATACAAATTAAGATG 4015
Qy      8522 TAGCGGTAGCCGCTGTGCATGCAAGAGCAAAACAAAGCAATTTACGAGATCAACAACT 8581
Db      4016 TTGCGAAAGATGATTTAGCAACAAAGCAAGGAACAAAGCACTTATTTGCAACAACTG 4075
Qy      8582 TAACTTTCAACCTGTAATGAGAGAAAGTATCTCAACTTGTGTCAAAGCTTTGGCTA 8641
Db      4076 CAGATGCACTACTGAGAAAAAAGAACAAAGCAAAATCAAC-----AAGTATGATGAGAA 4129

```


Db 3179 CTGATGCAAAATGTTAAACCAACAGCGAAACAGCAATGCTGATAAAGTACAGCGCAAG 3238
 Qy 7748 TAGGAAAAAATATGATGTTA---TTGTTAAAGATTAATCGAATTTGAGAGCGGAGCA 7804
 Db 3239 AAACGCAATTATGCTAATAACGGTGCACACAGAAAGAAAAACAGCTGGAAACAC 3298
 Qy 7805 AAGATTAACCGTAGAGCGGTAGCTGCCGAGCACTTATCTCAAAAGCAAAAGATGAA 7864
 Db 3239 AAGTTCAAACTGAAAAAACAAGCTGATACAGCAATTTGATGTCACATCAATTCAG 3358
 Qy 7865 TGAATTCAGAGTTGAAATTTGAAAGAGTATTTCAATGAGAAAAATGAGTAAGTACC 7924
 Db 3359 AAGTTGAAGCGGCTAAAAATGCGAAATTTGTAATGAAACCAATTCAGCCAGCAACA 3418
 Qy 7925 CTTTAAAGGAATG---GAGAGAAATCAATGTCAAAGTGGAAAAAGAAACAGAGTGA 7981
 Db 3419 CAACCTAAAGATATGCGAAACAGCAATTTGCTACAAAGCGAATGAACTTAAACAGCA 3478
 Qy 7982 CTGCTGAATCTCAGAGAGCTTCTGTAGAGCAGTAGCAGGCGAGAAATTTTCCGAA 8041
 Db 3479 TGGCTCAACGCAAGACATTTACTGCTGAAGAAATTTGCGCGCAATGGAAT---GTAG 3535
 Qy 8042 CAAAAGATGCCGAAAGCTTTATTTGAAAGTTAGTACAAATCCGAAAGATATTTTC 8101
 Db 3536 ATAAATGCTGTGACACAGCAAAATCAACATTTGAAGCTGTAATGATCAAAATGATG 3595
 Qy 8102 ATGCAATTAATGTAATGAGAAACACATAAATGAAAGTAACAGAGTTTCTAAG 8161
 Db 3596 ACCAAGCGAAACCACTGTTGAGGCTAGTATTTGATCAAGTAAACCAACAGCAATTA 3655
 Qy 8162 CAGTAACAGAGTTCTGTATTTGAGAGAGTTGAGTCAACCAAGCAGAAAGTACTGCTG 8221
 Db 3656 AAGCAACAGCAATTCAGACGGGAAAAATATATTTACAGCTCACTGATGTAATG 3715
 Qy 8222 GTAAATCTATGTTAGAGTTGAGAAAGAAATTTGTCAGAACAAATGATGAAATGCA 8281
 Db 3716 TAGATACAGCGAAAGATGCAAGTTAAATTCGATTCAAAGTACACACAGCAACAG 3775
 Qy 8282 TTTCTAAAGTAAAGTTGATGTAAGTAAGTAAGTAACTGCTAATCTTCTGATGATCAG 8341
 Db 3776 TTTAAATCAAAATCTAAAAATGATGTTGATCAAGCTGTGCAACTCAAAATCAAG 3835
 Qy 8342 GAATGAGAGAAATTTGCCGAGCAGAGTGAATCTTCTACAGACAAATGATTAATCTG 8401
 Db 3836 ATAAATCAACTGTGTGCTAACAATGAAAGAAAAATGCGCAAAATGATTTAAAG 3895
 Qy 8402 AATCGTAGTTGTTTACGAAACCAAGATTATGAAATATGATTAACAAAAATATA 8461
 Db 3896 CTAAAGAAAAACGTAATCAAGATATCTTAAATGCAAAACAACTAACATGTTACG 3955
 Qy 8462 TTTCAAGAGTCAATGCTTGTCTTTAAATGATACAAAGAAAGAAATAGATCTT 8521
 Db 3956 TTTAAAGATCAAGCAGTGTCTGATGTTCAAGGTAATTTACTGCAATCAACAAAT 4015
 Qy 8522 TAGCGGTAGCCGATGTGATGATGCAAGAAACAAAGCAATTTAGAGATCAAAAGT 8581
 Db 4016 TTGCGAAAGATGAATTTAGCAACAAAGCAAGGAAACAAACCTTATTTGCAAACTG 4075
 Qy 8582 TAACCTTACAACTGTAAATGAGAAACGTAATCTCAACTTGTGTCGCAAAAGCTTT 8641
 Db 4076 CAGATGCGACTACTGAAAGAAAAAGAACAGCAATCAAC-----AAGTATGCA 4129
 Qy 8642 AAAATGAAAAATTAAGAAATGTAAGAAAGAACTGAGAGACCTTATGCGAGCGAAAC 8701
 Db 4130 TAACGCAAGTATCAAAATATTTGAAATGACAGTCAATCATATGTAATCACTGCA 4189
 Qy 8702 CAGCGTTGAAATTTACAAAGAGTACTACAGAGCATTTGGTTCAGAGAAATTTGGGAA 8761
 Db 4190 AAGATTAATGCAATTTCAAGCAATTTGACCAATTTCAAGATCAACAGATGTTAAAG 4249
 Qy 8762 TTGAGATTAATTAAGAAAGATTTGCAAGATTAATGATTTGAAGTCAACGAGAGC 8821

Db 4250 CAAGAGCGAATTTGCTAACTGAATGCAAAATTAATTAATTAATTAATTAATG 4309
 Qy 8822 GAACCAAGAGAGCTTGTGGAAGAAATGATTTTCTGTGAATAATCAATTTCAAGCG 8881
 Db 4310 AGACTTAATGAAAGAAAGGTAACGATTTGACCAAGTATGAGCTGCAATTAAGAG 4369
 Qy 8882 AAAAATTCATCCATTGAAATGAATGAAGCAGAAATTTGTAACCGAAGTAAATGAG 8941
 Db 4370 GTTTAAATTAATTAATGACCAACTACTACAGGATGTAATTAATCTGCTAAAGTAC 4429
 Qy 8942 ATGCTTGAATGAATCTTGAATGATGATCTACAGGAAAAAGTGTGCTATGTTG 9001
 Db 4430 CAGTACAAAAAGTTCAACCACTTCAATGCTGTTTGAAGCCAGAGTAAATCTG 4489
 Qy 9002 GTATTGAAATGTTGATGTAATTAATGATTAAGAAAAATGTAAGCCAAATTCGAA 9061
 Db 4490 CTTTATGATCAAGCTGACGCTGATTAAGAAACAAATTAAGAACCAACCAATTCG 4549
 Qy 9062 GACATGCTATTGTGAAACCTACGAGAAACAAAGATTCAGCAATTTACAGAGCAAA 9121
 Db 4550 AACCAAGAAATTAAGG---TGCAAAACAGAAAGTTGATGATGAAATTAATCAAGG 4606
 Qy 9122 TAAATATTCTTGAAGAAAGAGACGCTGCACTGCACTGCAATTCGAATGACATTT 9181
 Db 4607 CAAATATTGACCAATCTTCAACAGATGAATGTTGATTAATGCAAGTTAAAGAG 4666
 Qy 9182 CCAATGAGATGATTAATTAATTAATTTGCAAGAGATGATGATCTTCTCAATTAAT 9241
 Db 4667 CTTAAATTAATGACGTTTAAACATTTAGTGTGATCAAAAAAGATGCTTTACTTA 4726
 Qy 9242 AAAATTCAAAAATTAATTAATTTACTTGAAGTC-----AAGTATGATGATG 9295
 Db 4727 AAGCTGATTAATTAATGCTAAATTAACGAGAGGATTAATCTTAAGCATGACTTCA 4786
 Qy 9296 TTCAATGAGGCTGCTGAGCAAGAGTGCAGAGCCAAAGCAGCTTATGTTAAAGAT 9355
 Db 4787 AAATGCGAAGCGAAACAAAACTGCTGATTTGAAACCAACGCGAGTCAAAATG 4846
 Qy 9356 AAATTAATTAATGAATTAATTAATTTAGAGGAAAAATTAACAGAGGAAACAT 9415
 Db 4847 ATCAAGCTACTCTTAAGATACATTAAGTTCAATTAATTAATTAATTAATTAAT 4906
 Qy 9416 ATGATATGCGGATATGATTAATTAATTAATTAATTAATTAATTAATTAATG 9475
 Db 4907 ACGATTAACAAATTCAGAGTAAAGAAAGATGATGATCAACAGATTAATTAATG 4966
 Qy 9476 CGATGCCAAAGATGATGCTGAGCTGCTGCGCACTGCACTATTAATTAATTA 9535
 Db 4967 CAGATCAGAAAGAAATTAATTAATTTCACTGACCTTAATGCAACCAAGATGA 5026
 Qy 9536 TAAATTAATTAATGCGATCGAGATTAATTAATTAATCTGCAAG 9581
 Db 5027 AAGCAATTAACAGAGTGAACCAAAATGTTCAAACTGCTTAAGAAAG 5072

RESULT 26
 US-10-972-13288/c
 ; Sequence 13288, Application US/10972079
 ; Publication No. US2005015317A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; APPLICANT: DENISE, Sue K.
 ; APPLICANT: ROSENFELD, David
 ; APPLICANT: KERR, Richard
 ; APPLICANT: BATES, Stephen
 ; APPLICANT: HOLM, Tom
 ; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
 ; FILE REFERENCE: MM1110-2
 ; CURRENT APPLICATION NUMBER: US/10/972,079
 ; PRIOR APPLICATION NUMBER: US 60/514,333
 ; PRIOR FILING DATE: 2003-10-24

```

; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13288
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894217622_1
US-10-972-079-13288

Query Match
Best Local Similarity 49.5%; Score 73.2; DB 9; Length 600;
Matches 186; Conservative 1; Mismatches 189; Indels 0; Gaps 0;

QY 7105 GTACGAGAAAAGTAAATGTAACGCAAAAGAGAAAATAATTAAAGCAACAGACGA 7164
DB 593 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 534
QY 7165 AATGACGAATCGAGACAGCAATCGAGCCATGCTGTGTAATAATTTGGACA 7224
DB 533 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 474
QY 7225 GCTGTGAAGATGAAAAAATTTGAGAGAAAGAACGAAGTTTAAACTTTAGAC 7284
DB 473 GAAGAAAGAAAAAGAAAAAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 414
QY 7285 GAAGTTAACAAAGAACAGATTAATAATGATGCTACGAAAAAATCTTACATCA 7344
DB 413 GAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 354
QY 7345 GCAGTATTTCTACAGAGTACTTCTGTAAAGCCGATGAGAGATCTCAGAGAA 7404
DB 353 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 294
QY 7405 GGAATTAAGCCATTGTGAAGCTTGTATATTGAAAAAATGTAGATTATTAACA 7464
DB 293 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 234
QY 7465 GAGACAGAAATTAATA 7480
DB 233 AAGAAAAAGAAAAAGAA 218

RESULT 27
US-10-282-122A-7578
; Sequence 7578, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreytch, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EPIRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7578
; LENGTH: 7446
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-282-122A-7578

Query Match
Best Local Similarity 41.8%; Score 72.4; DB 7; Length 7446;
Matches 804; Conservative 0; Mismatches 1091; Indels 27; Gaps 5;

QY 3782 TTCAAGTAAACAAAGTCTTGTTAAATATGATATTCGAATTCGAATTAATTAA 3841
DB 3608 TTAATTAACCACTGACAGTAATGAATACAGCAATTTAAATAACAATTGCAAG 3667
QY 3842 AAGTAAATCTTGAGTGAAGAACTCAAGTAGCAGCAGACAGGTTGAAAGCACTTA 3901
DB 3668 AATTCAGCTACACAGATGCAAGATGAAGAAAAACAGCAGCTGATGAGAAAGCA 3727
QY 3902 AAGAAAGTGAGACAAAGAAAAAGTTATCTATTGGAACTTCTGCTTATCAACTAG 3961
DB 3728 ATATCGAAATATGTAAAGCAATCAAGCAATTCAGCAGCAACTTAACGACAAAGTG 3787
QY 3962 TGAACATGAATTTCTGCAAAATCAGAAATATATACAGTACAGAGAAATCTGAAAGCC 4021
DB 3788 ATGAAGCGAAGCTATGCAAGAGCAAGCAATTAATGCGTTAACCAAAAGTTTAAAG 3847
QY 4022 AAAAAATGATGTTGATGCTCACTGCTTATCAAGCGGACCCAGTAGAGAGCTTTAA 4081
DB 3848 AACAGCGGCTAAAGATGATGATTCATCAAGCAAGCAAGCAAGTATTAATCAATA 3907
QY 4082 ATTTCAAGCTGGAAGTCAAAATGGAAGCTGAGGGGCTACTGATGCTGTCCTCAATTAA 4141
DB 3908 ATGATTAAGAACTTCAAAATGAAGAAAGAGAGCTTTTCAAAATTAAGCAAGCAG 3967
QY 4142 ACAACAAAGTAATGCTTCTATTAAGTGTGGAGATATCTAACGTTAAATCGACGAGC 4201
DB 3968 TTACAGAGCGGAAAAATATATTAAGCTGCAACTGATATATGATATGATATACAGCG 4027
QY 4202 CAAAAGCTTTTACGACCACTCAAGTGAAGTGTGCACTGACGACGAGGAGCAATTAA 4261
DB 4028 AAGAGCTGGAAGAAATTCATTAAGCAACCAACCAAGCAGCAG-----TTA 4078
QY 4262 GTTCTGAGACGGGATTAAGAAATTAATCAAGGGGCTTTCTGTCATAAG---ATTGACA 4318
DB 4079 AATCGAATCGAAAAATGAAGTTGATCAAGCTGTGCAACTCAAAATCAAGCAATTAATA 4138
QY 4319 ATGACGTGAGAGCTGATGATTAATCTTCATCGAAGAGCTAATGAATCAATATGTC 4378
DB 4139 ATACAACTGCGCTTCACTGAAGAGAAAAATGCAAGAAAGTTAGTTTAAAGCTTA 4198
QY 4379 TTGCAAGATGTCAAAGAGTCTGATCTAGCAAAAGAAATTCAGGCTTTAATGAATG 4438
DB 4199 AAGAAAAACGATCAAGATATCTTAAAGCAAAACAACTAATGATGTTACCAATTA 4258
QY 4439 GAAAAATTAATAATTTTAAGAAAGATCGGTATTAATGCACTGGAATGTTATTAATA 4498
DB 4259 AAGATCAACAGTGTGATTAATTAAGATTAATCTGCAATACAAATTAAGATGTTG 4318
QY 4499 CGAAGCAACTAGAAAAAGCAAAAGAAAAAGAGAGAGCGGCTCATTTGAATGCTGCTT 4558
DB 4319 CGAAGATGAATTAGCAACAAAAGCAAGCAAGCAAGCAAGCGGCTTATTGCAAACTGCA 4378
```

Oy	4559	TATGGTTGCTGGAAACGGATAAAATCCGGTGGAGAGTAGCTTAATTGACGTCAATCTGTTA	4618
Db	4379	ATGGCACTACTGGAGAAAAAGAAACAAGAAATCAACAAGATGACGCAATTAACACAAAG	4438
Oy	4619	AAAAATAAATTTAAAGCAGAAATTTGATGGAGCAATTAAGAAACCGGAGAGGATAAATTC	4678
Db	4439	GTAATCAAAAATATTGAAAAATGCACAGTCAATCGATGATGTAAACATCGCAAAAGATATAG	4498
Oy	4679	ATGGCAAAACATGTAAATGTGAGAGCAAAATCATCTACGTGTGTGTGAATCGGGCTCTG	4738
Db	4499	CAATTCAGCAATTGACCCCAATTCAAGCATCAACGATGTTTAAACGAATGCCAAGAC-A	4557
Oy	4739	GACTTGCTATGAGCAAAAGATGCTTTTTCAGAAATGGGATCTGAGCATGCGCAAGCTTAT	4798
Db	4558	GAATTGCTAACTGAAATGCAAAATTAATACTGAATATCTTAATTAATAGAGACT---	4614
Oy	4799	CAAAATGACACGATTCGAAAGGTGATAAAGAAAGAAATTTGCTGATTTCTTTAAATGTGA	4858
Db	4615	-----ACTAAAGAAAGAAAAAGGTAAACGATTTTGGACACGATTAGACAGCATATAGAG	4666
Oy	4859	AAGCAATTAATTCATTTCTTGAGGGTGAATGTGGGGAGAACCATTTGCCGCTCTCTTCTA	4918
Db	4667	AAGCTTTAAATTAATTAATTAATGACAGCACTTACAGGTGATGATATCTACTGCTAAAGATA	4726
Oy	4919	CGCGGTGAGAGCTGCTTTTGGATATACTCTTCATTAATTAACCTCTGCTTTGATTA	4978
Db	4727	CAGCAGTACAAAAAGTTCAACAACTTCATGCAAAATCTGTTAAGAAACAGCAGGTAA	4786
Oy	4979	CAGGACGAAAGTAAATCTCTTTAGTGGAAAGAAATCAAAAGTCATATGTACAGCTTGA	5038
Db	4787	AAGAATTAGATCAAGCTGCACTGATTAAGAAAAACAATATGAAACAAACCAATATCAT	4846
Oy	5039	ATGATTTTCATATTACAAACGTTTCTGCTGAGGCGCTCAAGATTAATAGCAGGCTGAA	5098
Db	4847	CACAACAAGAAATTAATGATGTGCAAAACAAGAAAGTTGATATCTGAATTAATCAACGAAA	4906
Oy	5099	TCGAGGAATGATCTGTCAATCGTGGTTCTGATGAACGAAAGCTTATGATGATTT	5158
Db	4907	CAAAATGTGATCAATCATCAACAAATGATATGTTGATTAATGACGATTAAAGAAAGAAAG	4966
Oy	5159	CTGAGTTTGAAGAGATAGTTCTTTCAATGTAGATGCAAAAGATCAAAAAACAATTAATA	5218
Db	4967	CTAAATTTAATGCAAGTTTAAACATTTAGTGTACAAAAGAAAGCTTTAGCTTAAATTTG	5026
Oy	5219	CAATTCGCGAAATGCAAAATGAGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCAT	5278
Db	5027	AAGATGCAATATAATGCTAAAGTAAACGAAGCGAGTAACTTAAGCATTCGACTTCAAGTG	5086
Oy	5279	CAAAATTTTGGAAAAACAATCAGTTATAGCTATTGTTAAAAACAATTAATTAACAACGCGGA	5338
Db	5087	AAATTTGCTGAAGCGAAACAAAACTGTGTAATTAACAAACGCGGATCAAAATGTTA	5146
Oy	5339	ATGATCAAGATAGAAAAAATATCATATGCACTGCAAAAGATTAATATATATGCAATACTA	5398
Db	5147	ATCAAGCTAATCTTAAGATGACATGTAAGTTCAAAATTCATTAATGACTTGATATAATTTA	5206
Oy	5399	TAGCAGTGGAGTTGAGAGACAAAGAGCCTCTGTGCAAGAGCTTCTGCAAGTACTA	5458
Db	5207	ACGATTAACAAATTTCCAAACAGGTAAAAAAGATCAGTCAACAGATTTATATCTTATG	5266
Oy	5459	CCTTGAATTAAGACAGTTTCTTCTCTCATGTTTGAATCAACATGATATTGACAAACATTTAGAG	5518
Db	5267	CAGATCAGAAAGAAAAATTAATATTTCAGCTGACATATGCAACACAGATBAAAAGCAAC	5326
Oy	5519	AAGAAATTAATGAAAAATAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATACGATC	5578
Db	5327	AAGCAATTAACCAAGTTGACCAAAATGTTCA---AATGCAATTTGAAAAATTAATATATG	5383
Oy	5579	AAGTGATCAAAATGCGACAGTGTCTTTCCGAGCAAGTGGACAAAGCTGCAATAGAGCTG	5638
Db	5384	GTTGTGATTAATGTGACGTTGATGATGATTAACCAAGAGTAAAGCAAGCAATTAATACTA	5443

Oy 5639 GAGTACGCGTTAAATTAATAATTACACAATACTTCGTGCACATATAAAAAATGTACTCAA 5638
| | | | |
Db 5444 TTCAAGTAGAGTACTACTGTAAACTTAAGGAACCAAGCATTTGGAGCTAAGAAGAA 5503

Oy 5639 AT 5700
| |
Db 5504 AT 5505

```

RESULT 28
US-10-425-115-132399
/ Sequence 132399, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 132399
/ LENGTH: 1150
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(1130)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_52232C.1
US-10-425-115-132399

```

Query Match	0.7%;	Score 70.6;	DB 8;	Length 1130;
Best Local Similarity	44.4%;	Pred. No. 0.041;		
Matches 366;	Conservative	0;	Mismatches 455;	Indels 3; Gaps 2

QY	8993	ATTAAGCCAGAAATTGGTGAACCCGGAAGTGTAAATGTGAATGCTTTGAATGAACCTGATG	8966
Db	272	ATTAAAAACCGATTAAATTATATGCGAGAGATTAAAGCTTTGATGAATGGAAGGCACTGAGGA	331
QY	8963	TAGATCTACAGAGAAAAAGTGCGCTATGCGTGAATTGCTATTGGAATGTGTA	9022
Db	332	TTGATTATTGCAAAAGACGTTGGTAAAGAAATGATGAAAGAAAAGAAACAATGGAA	391
QY	9023	ATATATGTGATTAAAGAAAAATGTAGAAGCCAAAATCGGAAGCAATGCTATTGTGAACCTA	9088
Db	392	TTGAAAGATGACACACAAACAAAGAAAAAAAAGAGCCAAACAATATACATTAATAA	451
QY	9083	CTGGAACAAAGAAATATCAAGCATTTACAGAGCAAAAATTAATTCTTGAAAAAGAG	9142
Db	452	AAAAACTAGACAGATTAAAAAAATATTAAACGAAAAACACTATCAAAAAAAAAGAAAA	511
QY	9143	ACGCTGCAAGCTGCACTGCAATTTCGAATGTACACATTTCCAAATGAGATGATATTA	9204
Db	512	AACATCAAAAAAGAAAAAACACAAACACACTGTACATCATTAACAATTAACCAAAATACAA	571
QY	9203	ATTGGCAAGCAAGTATGCACTCTCTCAATTAATTAACCAAAAAATTCAAAAAATATATTA	9267
Db	572	AAGCGAAAAATATATAAAAAATAAAAAATTAACAACATTAATAAAAAATATATTA	631
QY	9263	CTTTAGCACTCAATGATGCAATCGAATGTGAATGTTCAATGGGTGGCTGAACCAAGGTG	9322
Db	632	GAATAATTTAAGAAAAAT	691
QY	9323	CAGAGGCCAAAGCAGCAGTTAGTGTAAATGAATCAAACTAATATAGAACTATATATGTTGAT	9388
Db	692	TATATAAAAAT	751
QY	9383	TAGCAGAAAAATTTAAACAGAGGAAACATCAATGTATATATGCGGATATGATAAAAAT	9447

Db 687 AGAAAAGAGCAAGTAGAAAAAGAAAGAGCGCAAAAGCTGAAGAAACAGCAAGAGC 746
Qy 5805 AGTGAATAGATTTGTAATATACATAGCAGAAATTAATCATGCAAAAATCATGCGAA 5864
Db 747 AACAGAAAAAGACCAAAAGCTGAAGAAACAAAGTAGAAGAAAGAAAGAGCCGAC 806
Qy 5865 GGGAAATGTCGAGTTATTACAGAGTCTGATCGCGTAATTGCTAATTATGACAGAAAGT 5924
Db 807 AGCTGAAGAAACAAACAAAGCAGCAACAGAAACACACAGAGCTGAAGAACACAAATGC 866
Qy 5925 GCTGTAGAGCGCCGCGAGCAATAGAGACCTCAACAGTGTGAATGAATACAGATC 5984
Db 867 TAGAGCAATTAATAGCTCAACCTCAGAAACTGAGAGAACTCAAGTTGATGATACAGT 926
Qy 5985 TACAAAGCATATGTAAGAAATTTCTACAGTGAATGCTAAGAAAGAAACAGATGATTAAT 6044
Db 927 TCTTAAGATTTATATTAAGAAATCTGAAGTTACAGAAAGAAAGAAAGCTGAATGAAA 986
Qy 6045 TACTACTCAAGGCAAGTAGATTAAGTGTAGATAAAGTATTCAAAATCTTAATATTAA 6104
Db 987 AGTATTACCAAAAGATATTCTCAAACTTATCTAATGAAGAAATTAAGAAATAGCTTAA 1046
Qy 6105 CCAAGACTTATCACAAAAAAGAAATAGTAATTAAGAAAGATTTGTTCCAAATAGTTC 6164
Db 1047 TGAAGTACTTAAGAAAGCTTAACAAAGAAACGCAACAGCAAGCAACCTCCGTTTC 1106
Qy 6165 AGCTACTCATCTTAAATCTTTATTTGCAATGCGCTGCTGAGCAAGCCGAGT 6224
Db 1107 ACTAAGCAGCAATGCTGAAGAAACAAATTTAATCTATTACAGCAAGCTTAAGAGCAGC 1166
Qy 6225 GCGAGAACTGTTAATATCAACAGGTTATGAGAAACAGAAAGCTCTTTGAGA-AAAT 6283
Db 1167 TCACACAGACAGAGTTACTAAAAAGAAAGAACTGTAACTTACTGCGCAGAGATATAT 1226
Qy 6284 CATATTAATGCAAAACATTTCTGTAAATCAGAGATTCAGAGATTCAT 6338
Db 1227 CATATAACTTATTAAGAAAGAAATTCCTTAATGAAGCAGCTAAGCTGATCAAT 1281

RESULT 31
US-10-282-122A-35506
Sequence 35506, Application US/10282122A
General Information:
APPLICANT: Wang, Liangshu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELIIRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIORITY FILING DATE: 2003-02-20
PRIORITY FILING DATE: 2000-03-21
PRIORITY FILING DATE: 2000-03-21
PRIORITY FILING DATE: 2000-05-23
PRIORITY FILING DATE: 2000-05-23
PRIORITY FILING DATE: 2000-07-27
PRIORITY FILING DATE: 2000-05-26
PRIORITY FILING DATE: 2000-09-06
PRIORITY FILING DATE: 2000-09-06
PRIORITY FILING DATE: 2000-09-09
PRIORITY FILING DATE: 2000-10-23
PRIORITY FILING DATE: 2000-10-23

Qy 687 AGAAAAGAGCAAGTAGAAAAAGAAAGAGCGCGCAAAAGCTGAAGAAACAGCAAGAGC 746
Qy 5805 AGTGAATAGATTTGTAATATACATAGCAGAAATTAATCATGCAAAAATCATGCGAA 5864
Db 747 AACAGAAAAAGACCAAAAGCTGAAGAAACAAAGTAGAAGAAAGAAAGAGCCGAC 806
Qy 5865 GGGAAATGTCGAGTTATTACAGAGTCTGATCGCGTAATTGCTAATTATGACAGAAAGT 5924
Db 807 AGCTGAAGAAACAAACAAAGCAGCAACAGAAACACACAGAGCTGAAGAACACAAATGC 866
Qy 5925 GCTGTAGAGCGCCGCGAGCAATAGAGACCTCAACAGTGTGAATGAATACAGATC 5984
Db 867 TAGAGCAATTAATAGCTCAACCTCAGAAACTGAGAGAACTCAAGTTGATGATACAGT 926
Qy 5985 TACAAAGCATATGTAAGAAATTTCTACAGTGAATGCTAAGAAAGAAACAGATGATTAAT 6044
Db 927 TCTTAAGATTTATATTAAGAAATCTGAAGTTACAGAAAGAAAGAAAGCTGAATGAAA 986
Qy 6045 TACTACTCAAGGCAAGTAGATTAAGTGTAGATAAAGTATTCAAAATCTTAATATTAA 6104
Db 987 AGTATTACCAAAAGATATTCTCAAACTTATCTAATGAAGAAATTAAGAAATAGCTTAA 1046
Qy 6105 CCAAGACTTATCACAAAAAAGAAATAGTAATTAAGAAAGATTTGTTCCAAATAGTTC 6164
Db 1047 TGAAGTACTTAAGAAAGCTTAACAAAGAAACGCAACAGCAAGCAACCTCCGTTTC 1106
Qy 6165 AGCTACTCATCTTAAATCTTTATTTGCAATGCGCTGCTGAGCAAGCCGAGT 6224
Db 1107 ACTAAGCAGCAATGCTGAAGAAACAAATTTAATCTATTACAGCAAGCTTAAGAGCAGC 1166
Qy 6225 GCGAGAACTGTTAATATCAACAGGTTATGAGAAACAGAAAGCTCTTTGAGA-AAAT 6283
Db 1167 TCACACAGACAGAGTTACTAAAAAGAAAGAACTGTAACTTACTGCGCAGAGATATAT 1226
Qy 6284 CATATTAATGCAAAACATTTCTGTAAATCAGAGATTCAGAGATTCAT 6338
Db 1227 CATATAACTTATTAAGAAAGAAATTCCTTAATGAAGCAGCTAAGCTGATCAAT 1281

Query Match 0.74; Score 68.2; DB 7; Length 4997;
Best Local Similarity 41.0%; Pred. No. 0.2;
Matches 933; Conservative 0; Mismatches 1318; Indels 24; Gaps 5;
US-10-282-122A-35506

Qy 5215 AATCAATTTGCGGAAATGCAATGAGAAAGCGGCTGAGATTGAGCAAGTTGCT 5274
Db 1969 AAAGCAGCAGCAAGAAAGCAATTTGATGATGACAGCACTGCTAAAAAGAAAGCAATTAAT 2028
Qy 5275 CATACAAATATTGAAAAAAATCAAGTTATAGCTATTGTAAGAAACAGTAATTAACACG 5334
Db 2029 AATAGCTAGATGCAACCAAGAAAGAAAGATGAGCAATAGCAAGTTGATGACAGT 2088
Qy 5335 GCGAATGATCAAGTGAAGAAATTTCAATGATGCTGCAAGAAATTAATCTATGACCAAT 5394
Db 2089 GTAACAGCTGCAAAACCAAGCAATTAACACAGCAACCAAAATGATTAAT--GTAGACCAAG 2146
Qy 5395 ACTATGAGAGTGGAGTTGAGAGAGCAAAAGAGCTCTGTGCAAGAGCTTTCGCAAGT 5454
Db 2147 AACAAATAGCGGTATCTTCACTATTAATGATTAACCAACAG--AAGTACTTAAGAAAGCA 2205
Qy 5455 ACTACCTTGAATGAAGACAGTTCTTCTGATGATGATCAAGTATTAATGACAAAGATTAA 5514
Db 2206 GCAGCAGAAAGAAAGCAATTTGATGATGAAGTCTGCTGAAGAAAGCAATTTGATCTGTT 2265
Qy 5515 GAGAAAGAAATTAATGAAATTAAGAAAGCAATTTGATTTCTGCTGCTGAAATTCG 5574
Db 2266 GCCGATGCTACAGATGAGAAAGAAACAGCAGCTAAAGTAAAGTTGATGCTGAAGCTAAC 2325
Qy 5575 AGTCAAGTGTCACAAAATGCGACAGTGTCTCCGAGCAAGTTGAGCAAGCTGAGTAGA 5634
Db 2326 AAAGCAAAAGCAGCAATTTGATCAAGCTACTCAAAACAGATGAGCAAGCCAAATCA 2385
Qy 5635 GCTGAGATGAGCAATTAATTAATTAACAAATTAATTAATTAATTAATTAATTAATTAAT 5694
Db 2386 TCTGTTGAACAACTATTGAAGAAATTAACAACTGATTAAGAAATTCGTGACAAATA 2445
Qy 5695 CAAATGTAAGAAATGCTTTGTTGTAAGAAAGCAATTCATCTATTAATTAATTAATTAAT 5754
Db 2446 CAAGCAATGATGATGATGAGCAAAATGCTAAAAAGCAAGCAATTAATTAATTAATTAAT 2505
Qy 5755 ATTGAGCTGAGATTTGAGAGCTGAGAGCTGAGAGTGAAGCAAGTTCTGAGCAAGTAAAG 5814
Db 2506 ACTCAAGAAAGAAAGCAGTGAAGCAAAAGTGAATGATGAGAGCTGCTAAAGCTTAA 2565
Qy 5815 ATTGTAATTAATGATGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5874
Db 2566 ACAGGATTTGATCAAGCAACAAAGATGATGATGATGATGATGATGATGATGATGATGATG 2625
Qy 5875 GAGATTATTACAGATCTGATGCGGTAAATTTGCTAATTAATGAGAAACAGTCTGAGATG 5934
Db 2626 AATTCATTTCAAGTAATTCAGCTGATTAATTTGTAAGAAAGCAGCAGCAAGAAAGCAATT 2685
Qy 5935 GCCGTGAGCAATTAAGAGCTCAACAGTGAATGAATTAATTAATTAATTAATTAATTAAT 5992
Db 2686 GATGATGAGCAAGCTTAAGAAACAGAAATTTGATCAACCAAGCTTACACAAAGAA 2745
Qy 5993 ----CATATGTAAGAAATTTCAAGTGAATGCTAAAGAAAGCAAGATTAATTAATTAAT 6048


```

Qy 9072 TGTAGAACTACTGAA----AACAGAAATATCAAGATTATACAGAGCAAAAGTAATA 9127
Db 2619 AAAAAAAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATCT 2560
Qy 9128 TTCTTGAAAAGAGAGCGTGCAGCTGCAGCTGCATATGCAATGATGACATTTCCATG 9187
Db 2559 TAAAAAAGAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATC 2500
Qy 9188 AGATGATATTAATAAATTTGGCAAGACATGATGATCTTCTCAATTAATACCAAAAT 9247
Db 2499 TAAAAATTAATAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATC 2440
Qy 9248 CAAAAATATATTAATCTTAGCATCAAGTAGTGAATGCAATGCAATGCTGCGTGG 9307
Db 2439 TAAAAAAGAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATC 2380
Qy 9308 CTGAAGCAAGAGTGCAGAGCAAGCAAGCAAGTGAATGCAATGCAATGCAATGCA 9367
Db 2379 AAAAAAATTAACATATAAAAATTAATAATTAATAATTAATAATTAATAATTAATA 2320
Qy 9368 CTAAATATGTTGATTTAGCAGAAAAATTAATAACAGAGGAAACATCAATGATATGCG 9427
Db 2319 AAAAAAATTAATAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATC 2260
Qy 9428 GATATGATTAATAATTAATAATTAATAAGTGAACAAATTTCAAGGCTATTCGCGAT 9487
Db 2259 AAAAAAATTAATAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATC 2200
Qy 9488 GTCAAGCTGCAGCTGCTGCGCACTGCACATTTGAAAAATGAAGTAAATTTT----- 9543
Db 2199 AAAAAAATTAACATCTAAAAATTAATAATTAATAATTAATAATTAATAATTAATA 2140
Qy 9544 --AATATGCGATCCGAGATTTAAAAATTAATCTGCAAGATTGGAAGGAAAGCTAATA 9601
Db 2139 AAAAAAATTAATAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATCTAAAAATC 2080
Qy 9602 AAAAAAATTAATAATCTAAAAATTAATAATTAATAATTAATAATTAATAATTAATA 2020
Db 2079 AAAAAAATTAATAATCTAAAAATTAATAATTAATAATTAATAATTAATAATTAATA 2020
Qy 9662 CTTCGAAATGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 9721
Db 2019 AATCTAAAAACATTAATAATAAAGCAAGCTCTAAAAAATTAATAAATTAATAA 1960
Qy 9722 AA 9723
Db 1959 AA 1958

```

RESULT 33

```

US-09-742-096-4
; Sequence 4, Application US/09742096
; Patent No. US2002015541A1
; GENERAL INFORMATION:
; APPLICANT: DRULHE, PIERRE
; APPLICANT: DAUBERSTES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773US01V
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 1891
; TYPE: DNA
; ORGANISM: P. falciparum
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (2) ..(1891)
; OTHER INFORMATION:
US-09-742-096-4

```

```

Query Match 0.7%; Score 64.8; DB 3; Length 1891;
Best Local Similarity 41.9%; Pred.No.0.53;
Matches 731; Conservative 0; Mismatches 997; Indels 18; Gaps 5;

```

```

Qy 5913 TGCAGAACTGTCGTGAGTGGCCCGTGCAGCAATAGGACCTCAACCACTGTGAAATGA 5972
Db 1 TACATTAATCTAAGAGTGTAGATTAATAATAATTAATAAAGCCGAAGATATTAAGCA 60
Qy 5973 AATTACAGATCTCAAAAGCATATGTAAGATTTCTACAGTATGCTTAAGAAAGAAC 6032
Db 61 AATATCTTAATTAAGTATATAGAAACCAAAAGAAATTTATTTGCAATTTATTTAA 120
Qy 6033 AGATGATTAATTAATCTACTCAAGGCAAGTGAATGAATGCTGATTAAGTATTAATA 6092
Db 121 TAAATTTGCAAAATTCAGAAAAACAAGAAAGTGTATCAGAAAAATGTACAAGTCA 180
Qy 6093 TCTTAATTAATCAAGACT--ATCAAAAAAGAAATTAAGTAAATTAATAAGATTT 6150
Db 181 TGAATTTTAATGAATTAATTAATGAATGATGATGATGATGATGATGATGATGATGAT 240
Qy 6151 GTTACCAATAGTCA---GCTACTCACTTTTAAATCTTTATTTGCAAAATGCGCTGG 6206
Db 241 TTGAGAGAAAGTCAAGTAAATGACGATATTTTAAATGTTAGTAAAGATGTTCAACA 300
Qy 6207 TTCAGACAGCCGAGTGCAGAGAACTGTTAATATCAACAAGTTTATGAGAAACGA 6266
Db 301 AGAACACACACATGTTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGA 360
Qy 6267 AGCTCTGTAGAAAAATCTAATTAATAATGCAAAACATTAATCTGTAATAATGAGGATTA 6326
Db 361 AAGGTAGAAAGAAATGTAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGA 420
Qy 6327 CACGAAATCAATCGAGATGAGTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 6386
Db 421 AAGGTGAAAGAAAGTATAGCTTCAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTG 480
Qy 6387 TTCTTGATACCAATATTAATAAAGAAATACCAAGCAAGAGTTGAA--AACTAC 6443
Db 481 AATGTAGCTCAACTGTTGAAGAAATGCTGATCTCAACTGTTGAAGAAATGTTGCT 540
Qy 6444 AATGTGATGAAGTTTCGAGAAAGAAAGTGAAGAAATTAACAGCAATTTAAGCAAGAT 6503
Db 541 AAGGTGTGAAAGTGTGCTCCAGAGTGTGAAGAAAGTGTGCTCCAGAGTGTGAAGA 600
Qy 6504 TTCTCTTTTGGAGTCGAGTCGACAGCCGGGTAGAGACCGAGAGTGCAGAGAACCT 6563
Db 601 AAGGTAGCTGAAATATGTTGAAGAAAGTGTGCTGAAATATTTGAAGAAATGTAAGCT 660
Qy 6564 TTCGTAAATCAATTTGAGAGAAAGCAAGAGTGAAGTGAAGAGCAAGATTTTGT 6623
Db 661 AAGGTGAAAGAAAGTGTGAGAAATGTTGAAGAAAGTGTGAGCTGAAATGTTGAAGA 720
Qy 6624 AAAAAAGCTGAGATTAACAGCAAAACGTTAAGTTCTGTTGCAATTTGAATTTGCCAG 6683
Db 721 AAGGTAGCTGAAATGTTGAAGAAAGTGTGAGAAATGTTGAAGAAAGTGTGAGCTGA 780
Qy 6684 CGAGTGTGCTCAAAAGAGCTGGAATTTGAGACAGCTGAGTGCAGTTTACCAAGATGAATC 6743
Db 781 AATGTGAAAGAAATGTAAGTGTGCAAGCTGTTGAAGAAAGTGTGCTCAAGCTGTTGA 840
Qy 6744 AACAAGAGCAAGAGTGAAGAAATTTTAAATTAATGATCTGAAACAAATGATGATAT 6803
Db 841 AATGTAGCTCAACTGTTGAAGAAAGTGTGCTCAACTGTTGAAGA--AATGTAGT 897
Qy 6804 AGCAGAAATGAGTAAATTAATGAGTATGAGTGTGTTGACCGAGAGTGTGAATTTCTGC 6863
Db 898 TCCAGTGTGAAGAAAGTGTAGCTCAAGTGTGAAGAAAGTGTAGCTGAAATGTTGA 957

```



```
Db      989 GTGATGCGAACTGAGCAAAATTAAGCAGATATTAATGATAGTAGAAGTAAAG 1048
Qy      7130 CAAAAGAGAAAAAATAATTAGCAAAACAGCAAAATGCGAGATCGAGAGAGAGCA 7189
Db      1049 AAAATGTAGTCGCCAATACTAGAAAAAGTAAAGAACTACACTTAAGTGTACTA 1108
Qy      7190 TCGAGCCCAATGCTTGGTAAATATTTTGGACAGCTGTAGAAAGTAAAAAATTCG 7249
Db      1109 CTTTAGTATATATTAGAGGAGATACAAAGAAATATCTATTAATGATCTATAGAG 1168
Qy      7250 AAGGAAAGGACAGAGATTTAAAACTTACAGAAAGTTACAAAGAACAGATAAA 7309
Db      1169 AAAAATTGAGAACTCCACGAAATGTATTAAGCCGCTTAGAAAAATACCCAAAGTG 1228
Qy      7310 AAGTAATGATGCTACGAAAAAACTTACATCAGCAGTATTTCTACAGAAAGTACTT 7369
Db      1229 AAGAGAAAAAGAAAGTAAATGATGTAATTGAAAGAAAGTAAAGAAAGTCCGTACCA 1288
Qy      7370 CTGTAAAGCCGATAGAGAGATCTCAGGAGAAAGAAATTTAAAGCCATTGTGAAGCTT 7429
Db      1289 CTTTAATGAAATCTGTGAACAGCAGAAAGAAAGTACAAATTACGAAATAT 1348
Qy      7430 CTGATATTATTGAAAAAATGTAG-----ATATTACACAGAGACAAAGATTAATATCA 7483
Db      1349 TTGAAAAATTTAGAAAGAAATGCAATGAGAAAGTAAAGTTGACAGAAATTTAGAGA 1408
Qy      7484 CTTTACGTGCTGTTTGGAACTGCAAGCTTCTCCGATCAGAAACAGTGGAGTTA 7543
Db      1409 AATTAAAGAAACGTGATTTTAACTGATTAAGTAAAGTAAAGAAACAGTAAAGTTA 1468
Qy      7544 CAATATTATTAAGAAATTCGGAGTTACTGTGAAAAATCTTTTGTGAAGCAGCTGAA 7603
Db      1469 GCGGAGAAAGTTTAAAGAAACATGAAATCGATTAAGCATTTTGTAGTAAATTTGATA 1528
Qy      7604 AAGTAATGTTAGATCGATATTACAGAAATGTTGCTTT 7643
Db      1529 ATGTAAGGAATACAGAAATTTATTACAGTATGTT 1568
```

RESULT 35

```
US-10-425-115-115706/c
; Sequence 115706, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 115706
; LENGTH: 1121
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1121)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_37006C.1
US-10-425-115-115706
```

Query Match 0.7%; Score 63.8; DB 8; Length 1121;

Best Local Similarity 44.8%; Pred. No. 0.63;

Matches 290; Conservative 0; Mismatches 353; Indels 5; Gaps 1;

Qy 7110 AGAAAAAGTAAATGTAACGCAAAAGAGAAAAAATATTAGCAAAACAGCAGCAAAATGC 7169

```
Db      948 AAAAAAGAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 889
Qy      7170 AGAATGCGAGAGACAGCAATTCGAGCCAAATGCTGTGTAATTAATTTGGACAGCTGT 7229
Db      888 AAAAAAGAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 829
Qy      7230 AGAAGATGAGAAAAATTCGTAAGGAAAAAGCAAGAGTTTAAAAAATTGACGAAAGT 7289
Db      828 AAAAAAGAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 769
Qy      7290 TACAAAGACACAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 7349
Db      768 AAAAAAGAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 709
Qy      7350 TATTTCTACAGAGATCTTCTGTAAAGCGAGATAGAGAGATCTCAGGAGAGAGAT 7409
Db      708 TAAATTAATAAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATA 649
Qy      7410 TAAAGCAATGTGTAAGACTTCTGATATTATTGAAAAAATGTAGATATTACACAGAGA 7469
Db      648 TAAAGTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 589
Qy      7470 CAAGATTAATATCACTTCTACTGCTGTGGAACTGACAGCTCTGCTCCGATCAGG 7529
Db      588 AATTAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 534
Qy      7530 AACAGTGCAGTTCAATATTAATAAAGAAATTCGGAGTTACTGTGAAAAATTCCTTTGT 7589
Db      533 AAAAAATTAATTAACAAAAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 474
Qy      7590 GAAAGCAGCTGAAAAAGTAAATGTTAGATCGATATTACAGAAATGTTGCTTTAACAGC 7649
Db      473 AAAAAACCAAAACATTAACAAACCTGATACAAATTAATAAATAAATAAATAAATAAATA 414
Qy      7650 ATATCAAGCTCTGTAGAGCATTTGGGAATGAGAGCTGCTATGCAATTAATTTCTAA 7709
Db      413 ATATCCATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 354
Qy      7710 TGGAAGATCAAAATATCGATTTAAAAATTTCAAGCTATTAGAAAAA 7757
Db      353 AACATTAACACACACTCAATTAATAATATCTACATAGTATTAATAA 306
```

RESULT 36

```
US-10-425-115-151933/c
; Sequence 151933, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 151933
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(996)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_70090C.1
US-10-425-115-151933
```

Query Match 0.6%; Score 63.2; DB 8; Length 996;

Best Local Similarity 44.0%; Pred. No. 0.76;

Matches 310; Conservative 0; Mismatches 394; Indels 1; Gaps 1;

Oy	363	TGAGGAAATTTATTTCTTAAGCTCGAAGGAGTGGAGTAAAGAAAAATGAGTTAT	422
Db	39818	TAATTAATTAATGACATTAATTCACACATTAATTTATTAATAATTAATTTGATTA	39759
Oy	423	CAATGCTGTTCTTTTCATTCATTAATTCGAAAAAGATGATTTTAAAGAGCTTTGGA	482
Db	39758	AAATTCATATTCAAAACATATTTATTAACACGTTAAATGTAATAAATAATCTTTTGA	39699
Oy	483	AGAAAGCAACATGCTAAAGTTTAAATGGAATCATTCAGTATGATGAAAAAGTAAAT	542
Db	39698	AAATTAATAATGATTAATTAATAGTAAATGCTTCATTAATATATGAGAAAAATATA	39639
Oy	543	TCCATTGAATCCGAATGGAAGCATTAACGTAGAAAGAAAAATCAATGCTTTGAAGCAT	602
Db	39638	AAATACATTTTATTTCTGCTCATTAATAATAGTTTCAATTAATGAAATGAAAAATCTGG	39579
Oy	603	CGGTTTATATGCGGCGGATTTAGATGAAAGATTAATGCAATCTAAAGACAGAAATTAAC	662
Db	39578	ATTTTTCATGCTTCGATGTGTATTAATAAATTAATAGAAAAATTAATTAATTTGGCTTAA	39519
Oy	663	AGATTTTAAATTTAGTCAATATTAAGTATGCAATTAATTCGTCTGACCGGAGATT	722
Db	39518	TAATTAATAAAACATTAATAATTAATAATTAATTAATTAATTAATTAATAATAATA	39459
Oy	723	AAAAAGTACCAAGCAAAATCTGGAGATTAATTTCTTTCAGCTCACATGAT-----TC	776
Db	39458	TAATTAATTAATTAATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA	39399
Oy	777	TCCCTCAAAAAGCTATGGGAAAAATTCGACTGTGGAAAGAGAAATAGAAATATGTAA	836
Db	39398	TCTACATCCATATTTATTTAAATCAATTTATTAAGATGGATTAATTTAAATTAATTA	39339
Oy	837	AGGAAATCCAAAGCAATATTTGAATCTG-ATGCTGATTTGGAAGCAGATGCAATATA	895
Db	39338	ATATTAATTTATTTGATTAATTAATTAATTAATTTAAATTAATCAATTTATCTATATATA	39279
Oy	896	AAATTAAGTGCMAAGCTACAAATGAGGATTTTAAGAAAGAGGGGAAAAAGAACTT	955
Db	39278	AAATTAATGAAATTAATTAATAAATAATTTAATTAATAATTAATGAAAAAACAATAATTA	39219
Oy	956	ATTAACATCTCTTTAAGTTATCAAGATGCGAAGCTCCGTAAGATTAATAAGAAAG	1015
Db	39218	ATTTAGCAAAATCAATTAATTAATCTATCTTAATGAAAAAACAAGAAATTAATTTAAAG	39159
Oy	1016	TCATGGAAGAAATGTTGACATTAACAGTGAAGCAAGAAATTTCTATGAT	1065
Db	39158	AAATTAATTAATGATGATTAATAATGTAATTAAGAAAGATTTTATTAAT	39109

Search completed: November 26, 2005, 06:34:14
Job time : 4553 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: November 25, 2005, 20:31:24 ; Search time 307 Seconds
(without alignments)
4708.018 Million cell updates/sec

Title: US-10-647-057-8

Perfect score: 9726
Sequence: 1 atgagcgagcacaataa.....gagaaagggagaaatga 9726

Scoring table:
Gapop 10.0 , Gapext 1.0

Searched: 3205263 seqs, 74304013 residues

Total number of hits satisfying chosen parameters: 6410526

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, NA, New.*
1: /cgn2_6/pdata/2/pubpna/US10_NEW_PUB.seq.*
2: /cgn2_6/pdata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/pdata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/pdata/2/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/pdata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/pdata/2/pubpna/PCT_NEW_PUB.seq.*
7: /cgn2_6/pdata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/pdata/2/pubpna/US11_NEW_PUB.seq3.*
9: /cgn2_6/pdata/2/pubpna/US60_NEW_PUB.seq.*
10: /cgn2_6/pdata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	79.2	0.8	4965	1	US-10-485-517-43 Sequence 43, Appl
C 2	63.2	0.6	2407	1	US-10-485-517-42 Sequence 42, Appl
C 3	51.4	0.5	3543	1	US-10-793-626-3493 Sequence 3493, Ap
C 4	50.2	0.5	2189	1	US-10-485-517-121 Sequence 121, App
C 5	50.2	0.5	3174	1	US-10-485-517-57 Sequence 57, Appl
C 6	48.4	0.5	3347	1	US-10-485-517-12 Sequence 12, Appl
C 7	47.8	0.5	3543	1	US-10-793-626-4149 Sequence 4149, Ap
C 8	47.8	0.5	3840	1	US-10-793-626-3187 Sequence 3187, Ap
C 9	45.6	0.5	1293	7	US-11-074-176-131 Sequence 131, App
C 10	45.6	0.5	2275	1	US-10-485-517-11 Sequence 11, Appl
C 11	45.6	0.5	3309	1	US-10-793-626-3637 Sequence 3637, Ap
C 12	45.4	0.5	663	1	US-10-793-626-1821 Sequence 1821, Ap
C 13	45.4	0.5	3180	1	US-10-793-626-3712 Sequence 3712, Ap
C 14	45.4	0.5	3650	1	US-10-793-626-3756 Sequence 3756, Ap
C 15	45.2	0.5	570	1	US-10-793-626-2229 Sequence 2229, Ap
C 16	45.2	0.5	2858	1	US-10-793-626-3777 Sequence 3777, Ap
C 17	45.2	0.5	3002	1	US-10-793-626-3968 Sequence 3968, Ap
C 18	45.2	0.5	3063	1	US-10-793-626-3432 Sequence 3432, Ap
C 19	45.2	0.5	3295	1	US-10-793-626-3676 Sequence 3676, Ap
C 20	45.2	0.5	3379	1	US-10-793-626-4138 Sequence 4138, Ap
C 21	45.2	0.5	3518	1	US-10-793-626-4114 Sequence 4114, Ap
C 22	45.2	0.5	3841	1	US-10-793-626-3497 Sequence 3497, Ap
C 23	45.2	0.5	3985	1	US-10-793-626-3856 Sequence 3856, Ap

24	45.2	0.5	4719	1	US-10-793-626-2905 Sequence 2905, Ap
25	45	0.5	2119	1	US-10-485-517-13 Sequence 13, Appl
C 26	45	0.5	2621	1	US-10-793-626-4448 Sequence 4448, Ap
C 27	45	0.5	3589	1	US-10-793-626-3353 Sequence 3353, Ap
C 28	44.8	0.5	1206	1	US-10-793-626-835 Sequence 835, App
C 29	44.8	0.5	3354	1	US-10-793-626-4188 Sequence 4188, Ap
C 30	44.8	0.5	3497	1	US-10-793-626-4025 Sequence 4025, Ap
C 31	44.6	0.5	3066	1	US-10-793-626-3818 Sequence 3818, Ap
C 32	44.6	0.5	3170	1	US-10-793-626-1849 Sequence 1849, Ap
C 33	44.6	0.5	4329	1	US-10-793-626-2051 Sequence 2051, Ap
C 34	44.4	0.5	1500	1	US-10-793-626-1557 Sequence 1557, Ap
C 35	44.4	0.5	3145	1	US-10-793-626-4198 Sequence 4198, Ap
C 36	44.4	0.5	3328	1	US-10-793-626-4195 Sequence 4195, Ap
C 37	44.4	0.5	3335	1	US-10-793-626-3503 Sequence 3503, Ap
C 38	44.4	0.5	3771	1	US-10-793-626-4066 Sequence 4066, Ap
C 39	44.2	0.5	2454	1	US-10-793-626-49 Sequence 49, Appl
C 40	44.2	0.5	2454	1	US-10-793-626-1527 Sequence 1527, Ap
C 41	44	0.5	1044	1	US-10-793-626-461 Sequence 461, App
C 42	44	0.5	1044	1	US-10-793-626-1267 Sequence 1267, Ap
C 43	44	0.5	3153	1	US-10-793-626-3348 Sequence 3348, Ap
C 44	44	0.5	3267	1	US-10-793-626-4042 Sequence 4042, Ap
C 45	44	0.5	3618	1	US-10-793-626-3564 Sequence 3564, Ap

ALIGNMENTS

RESULT 1	US-10-485-517-43/C	Application US/10485517
Sequence 43, Appl	US/10485517	
Publication No. US20050256299A1		
GENERAL INFORMATION:		
APPLICANT: University of Sheffield		
APPLICANT: Biosynex Incorporated		
APPLICANT: Foster, Simon		
APPLICANT: Mond, James		
TITLE OF INVENTION: Antigenic Polypeptides		
FILE REFERENCE: P100629MO		
CURRENT APPLICATION NUMBER: US/10/485,517		
PRIOR FILING DATE: 2004-02-02		
PRIOR APPLICATION NUMBER: GB 0118825.9		
PRIOR FILING DATE: 2001-08-02		
PRIOR APPLICATION NUMBER: GB 0200349.9		
PRIOR FILING DATE: 2002-01-09		
NUMBER OF SEQ ID NOS: 424		
SOFTWARE: Patencin version 3.1		
SEQ ID NO 43		
LENGTH: 4965		
TYPE: DNA		
ORGANISM: Staphylococcus aureus		
US-10-485-517-43		
Query Match	0.8%; Score 79.2; DB 1; Length 4965;	
Best Local Similarity	41.3%; Pred. No. 1.8e-05;	
Matches 998; Conservative	0; Mismatches 1378; Indels 42; Gaps 5;	
3782	TTCAAGTAAACAAAGTCTTTGTTAAATAGTATTTTGAATGCAATTAATTTA	3841
4395	TAAATTAACCACTGACGTAATGAATCAAGCACTTTTAAATTAACAAATTCAG	4336
3842	AAGTAATCTTGAAGTGAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	3901
4335	AAGTCACTGACGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	4276
3902	AAGAAAGTGAAGCAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	3961
4275	ATACGTAATGTAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	4216
3962	TGAACATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	4021
4215	ATGACACTGAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	4156
4022	AAAAATGATGTGTATGTCATGCTTATCAAGCGGACCAAGTGAAGAGAGCTTTAA	4081

Db 4155 AACAAAGCGCTAAAGATGAAATTGATTCATTTACAAGCAACGCAACAAATGTTATCAATA 4096
Qy 4082 ATTCAAGAGCTGAAAGTCGAAATGGAACTGTAGGGGCTACGTGACTGTGGCCAAATTAA 4141
Db 4095 ATGATCAGAACGCTACACACAGAAAGAAAGAACAGCTATTTCACAAATTAGCAACGCG 4036
Qy 4142 ACACAAAAGTAAATGCTTCTATTAGTGTGTGGAGATATCTAACGTTAATCGAGCGAG 4201
Db 4035 TTACAGACGCGAAAAATAATTACAGCTGCAACTGATGATATAGTGTATGAGCGGGA 3976
Qy 4202 CAAAAGCTCTTTAGCAACCACTCAAGTACTGTGCACTGACGACGAGGAGGACAAATTA 4261
Db 3975 AAGAGCGCTGGAAGAAATTCAAATTCAAAGCACCAACGACCAACAGCGGTTAAATCAATG 3916
Qy 4262 GTTCTGAGCGGGATTAGGAAATTATCAAGGGCTGTTCTGTCAATAAGATTGACAAATG 4321
Db 3915 CTAACAAATGATTTGATTCAGAGCTGTGACAACTCAAAATCAAGCAATTGATTAATACAACTG 3856
Qy 4332 ACGTGAAGCTAGCGTTGATTAATCTTCATCGAAAGAGCTAATGAAATCAATGTCATTG 4381
Db 3855 GTGCTACACTGAAAGAGAAATATGACGAAAGATTAGTTTAAAGCTTAAGAAAAAG 3796
Qy 4382 CCAAGATGTCAAAGAGTCTTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGA 4441
Db 3795 CGTATCAAGATATCTTAAATGCACAAAACAATATGATGTACGAAATTAAGATCAAG 3736
Qy 4442 AAGATTAATAATATTAGAAAGATCGTGATTAATACGACTGGAATGTTATTATACGA 4501
Db 3735 CAGTTGCTGATATTCAAGGTATTACTGCAATACAACTAAAGATGTTG-----CGA 3682
Qy 4502 AAGAACCACTAGAAAAAGCAAAAGAAAAAGAGAGCGGTCAATTGTAATGCTGCTTAT 4561
Db 3681 AAGATGAATTACGACAAAGAGAAAGCAAAAGAGCGCTTATTTGACAACTGACAGATG 3622
Qy 4562 CGGTGCTGGAACGATTAATCCGCTGAGAGAGTACTATTGCACTCAATCTGTTAAAA 4621
Db 3621 CCAGTACTGAAAGAAAAAGAAAGCAAGCAATCAAGTAGACGACATTAACACAAGTA 3562
Qy 4622 ATTAATTTAAAGCAGATTGAGTGGAGAGCAATTAAGAGCGGAGAGGATAAATTCAATG 4681
Db 3561 ATCAAAATATTGAAATATGCAAGTCAATGATGATTAACACTGCAAAAGATTAATGCA 3502
Qy 4682 CGAAACATGTAATGTGAGGCAAAATCATCTACTGTGTGTGTAATGCGGCTTGAGC 4741
Db 3501 TTCAAGCATATGACCAATTTCAAGCATCAAGATGTTAAAGATATGAAAGC-GGA 3443
Qy 4742 TTGCTATCAGCAAGATGCTTTTTCAGGAATGGGATCTGAGACATGCAAGCTTATCA 4801
Db 3442 TTGCTAATCGAAATGCAAAATTAATTACTGAATTAATTAATATGAGACT----- 3389
Qy 4802 ATGACACGATTGCCAAGGTGGAATTAAGAGAAATTTCTGCTGATTCCTTAATGTGAAG 4861
Db 3388 -----ACTATAGAAAGAAAAAGGTAAAGATTTGACCAAGTATAGAGCATATGAAG 3334
Qy 4862 CAAATTAATTCATTTCTTGGGTGAATGTGCGGAAACATTCGCGGTTCTCTTTCTAGCG 4921
Db 3333 GTTTAAATATATTAATGACAGCACTACTACAGGTATGTAACTACTGTGTTAAAGATCAG 3274
Qy 4922 CGGTAGAGCTGCTTTTTCGAAATTAATCTTCTCATTAATAAACTCTGCTTTGATTACG 4981
Db 3273 CAGTACAAAAGTTCAACAACATTCATGCAATCTGTGTTAAGAACAGAGGTAAAAAAG 3214
Qy 4982 GAAACAGGTAATCTTTTATGTGAAAGAAATACAAAGTCAATGTACAGCTTTGAATG 5041
Db 3213 AATTAGATCAAGCTGAGCTGATTAAGAAAAACAACAAATTAACAAACACAAATGATCAC 3154
Qy 5042 ATTCTCATATTACAAAGCTTTCTGTGAGAGCGCTGCAAGTATTAAGCAGGCTGGAATG 5101
Db 3153 AACAGAAATTAATGATGCAAAACAAAGAGTTGATCTGAATTAATTCAGGGAACCA 3094
Qy 5102 GAGGAATGATCTGTCAATCGTGTCTGATGAACGGAAGCTTATGATTGATTCTG 5161

Db 3093 ATGTCGATCAATCATCAACAAATGAATATGTTGATTAATGAGCTTAAGAGAAAAAGCTA 3034
Qy 5162 AGTTGAAGAGATAGTTCTTTCAATGTATGATGCAAAAGATCAAAAAACAATAATACAA 5221
Db 3033 AAATTAATGAGCTTAAACATTTAGTGTGATCAAAAAAGATGCTTTAGCTTAATTTGAAG 2974
Qy 5222 TTGCGGAATGCAAAATGAGAAAAAGCGCTGAGATTGAGCAACAGTTGCTCATACAA 5281
Db 2973 ATGCAATATATGCTTAAAGTAAACGAAAGCGATTAATCTTAAGCATTCGACTTCAAGTAAA 2914
Qy 5282 ATATTGAAAAACATCACTTATAGCTATTGTAAATAACAGTAAATTAACAACGGAATG 5341
Db 2913 TTGCTGAAGCGAAACAAAACTGTGTAATTAACAACTGCGGATCAAAATGTTAATC 2854
Qy 5342 ATCAAGATAGAAAAATAATCATGATGACTGCAAAAAGATTATACTATGACCAATACTATAG 5401
Db 2853 AAGCTACTTCTAAAGATACATTGAAGTTCAAAATTCAATATGACTTATATTAATTAACG 2794
Qy 5402 CAGTGGAGTGGAGAGCAAAAGAGCGCTGTGTGCAAGAGAGCTTGTCAAGTACTACT 5461
Db 2793 ATTACAAATTCACACAGTAAABAAAGATCAGCTACAAAGATTTATATGCTTATGCAAG 2734
Qy 5462 TGAATAGACGTTCTTCTCATGTTGATCAAACTGATATTTGACAAAGATTTAGAGGAAG 5521
Db 2733 ATCAGAAAGAAATTAATTTTCAGCTGACATTAATGCAACAAGATGAAAAAGCAACAA 2674
Qy 5522 AAAATTAATGAAATTAAGAAAAAGCAAAATGTTAATGTTCTGTGCTGAAATAACAGTCAAG 5581
Db 2673 CAATTAAGCAAGTTGACCAAAATG---TTCAAACTGCAATTGAAAGCTTAAATATGTTG 2617
Qy 5582 TGTGTCAAATGCGACAGTGGCTTTCCGAGAGAGTGAACAAGCTGACGTAGAGCTGAG 5641
Db 2616 TGGATTAATGTGAGCTGATGATGCAATTACAAAGGTAAGGACAGCAATTGATGCTATTTC 2557
Qy 5642 TAGCAGTTAATTAATTAACAAATACTTGTGCAATATTAATAAATAGTACTCAAAATG 5701
Db 2556 AAGTATGATGCTACTGTTTAACTTAAGCAACCAAGCTATTTGAAGTTAAACAGAGTA 2497
Qy 5702 TACGAATGCTTTGGTAAAG-----CAATCTCATTCATCTA 5740
Db 2496 CGAAAGATCTATTGATTAAGTGAACGAGTTAATCTGCTGAAGAAAACTGAAAGACTAG 2437
Qy 5741 TTAACCAATTTGAATTTGAGAGCTGAGAGTTGAGCTGAGAGCTGAGAGTGAACAGTTCTG 5800
Db 2436 CAATGATTAACCAATTAACAGATCAAGTAAACAAAGTATTAATGATGATCAACAACTG 2377
Qy 5801 TAGCAGTAATTAAGATTGTAATTAATACGATAGCAGAAATTAAATCATCAAAATCACTG 5860
Db 2376 CTGAAGTTGAAAAAGCGAAAGCTCAAGGACTTGAAGCAATTGATTAACATTCAAATCGACT 2317
Qy 5861 CGAAGGGAATGTGGAAGTTATTACAGAGTGTGATGCGGTAAATGCTAATTAATGACAGGA 5920
Db 2316 CAACGAAABAAACAAAAGCTATCGAAGATTTAGAACTGCACTAGACCAAGATTGAAGCAG 2257
Qy 5921 CAGTGTCTGAGAGTGGCCGCTGAGCAATAGAGCTCAACCAAGTGTGATGAATTAACAG 5980
Db 2256 GTGTAAATGTCAAGCTGATGCTACAACTGAAGAAAAAGAGCGTTTACGATGCTTAG 2197
Qy 5981 GATCTACAAAGCATATGTAAAGATTCTACAGTATGCTTAAGAGAAACAGATGATT 6040
Db 2196 AAGACATTTTAATCAAAAGCACTGAAGATATTTCTGATCAAACTACAAATCAGAAATCG 2137
Qy 6041 ATATTACTACTCAAGGCAAGTAGAATAAGTGTGATTAAGTATTCAAAAATCTTAATA 6100
Db 2136 CTACTGTCAAAATATGTGCGCTGTAACAATTAAGCAACAGATTTATCTCTGAAGTTA 2077
Qy 6101 TTAAGGAAGCTTTACAAAAAAAGAAAAATTAAGTATTAATAAAGATTTGTTACCAATA 6160
Db 2076 AGAAAAATGCTTTGGAACCAATCAGAGAAAGTGTTAACACACAAATTAATAATTAATAA 2017
Qy 6161 GTTACGCTACTCATCTT 6178
Db 2016 ATGCAGATGCAAGATGCAT 1999

RESULT 2

US-10-485-517-42/c
; Sequence 42, Application US/10485517
; Publication No. US2005025629A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynex Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825, 9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349, 9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-42

Query Match 0.6%; Score 63.2; DB 1; Length 2407;
Best Local Similarity 43.3%; Pred. No. 0.0089;

Matches 296; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 5185 AATGTAGATGCAAAAGATCAAAAACATTAATACATTTGCCGAAATGCAATGAGGA 5244
DB 874 AACGCATATATATACAGTAAACAGAAATTCAAAATAGCAATCTTCACTACAGAGAA 815
QY 5245 AAGCGCGTGGAGTGGAGCAAGTTCATCAAAATATTGAAAAACATGATTATA 5304
DB 814 AAACAAGCTGCAATACAGATTTAGATCTAAAAGACAGAACAGACAAATCTTGAT 755
QY 5305 GCTATTGTAAAAACAGTAAATTTACACCGCAATGATCAAGATTAAGAAAAATATCAAT 5364
DB 754 GCTGCAAAATCAAAACAGTATGTAACAAGCTTAAGACAAATGATTTGCTGCAATTAAT 695
QY 5365 GTGACTGCAAAAGATTATCTATGACCAATATCTATGACAGTGGAGTGGAGCAAAA 5424
DB 694 CAAGTTCAGCTGCCCACTAAGAAATCGATGCAAAACGGAAATCGCTCAAAAAGCA 635
QY 5425 GAGGCTCTGTGCAAGAGCTTCTGCAAGTACCTGTAATAAGACAGTTTCTTCAT 5484
DB 634 AGTGAACGTAAAAACAGCAATTTGAAGCAATGATTCGACTGAAAGAACAAAGCA 575
QY 5485 GTTGATCAAACTGATATTGACAAAGATTTAGAGAAAGAAATATGAAATTAAGAAAG 5544
DB 574 GCGAAAGACAAAGTGGATCAAGCAGTAGTACTGCAAAAGCTGATTAATGATTAATGCTGCA 515
QY 5545 GCAAAATGTTATGTTTCTAGCTGAAAATACAGAGTCAAGTGTCAAAAATGCGACAGTCTT 5604
DB 514 GCAAAACATGATGTGATTAATGCAAAAACATAAATGAACTCAATCGCAGCATTAACA 455
QY 5605 TTCGGAGCAAGTGGACAGAGTGTGAGAGCTGGAGTACAGTTAATAAATTAACAA 5664
DB 454 CTGATGCAAAATGTTAAACAGCAGCAAAACAGCAATTCAGATTAAGTAAAGTCAAGCTCAA 395
QY 5665 AATACCTCTGACATATATAAAATAGTACTCAAAATGTACGAAATGCTTTGTTAAAGC 5724
DB 394 GAAACAGCAATGATGAAATTAACGCTCAACAACTAAGAAAGAGAGCTGCTAAACAA 335
QY 5725 AATTCATTCATCTATTTAAACCAATTGGAATTGAGAGTGGAGTGGAGCTGAGAGCT 5784
DB 334 CAAGTTCAAACTGAAAAACACAGCTGATGCGCAATAGATGACAGACATACAAATGCG 275
QY 5785 GAGTGAAGGTTCTGAGCAGTGAATTAATTTTAATAATCGATAGCAGATTAAT 5844

RESULT 3

US-10-793-626-3493
; Sequence 3493, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3493
; LENGTH: 3543
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3493

Query Match 0.5%; Score 51.4; DB 1; Length 3543;
Best Local Similarity 42.8%; Pred. No. 1.2; Indels 9; Gaps 1;

Matches 329; Conservative 0; Mismatches 431; Indels 9; Gaps 1;

QY 5141 AAGCTTAACTAGTATCTGAGTTTGAAGAGTAAGTCTTCAATGATGACAAAAG 5200
DB 601 AAGCAATCAGTGGAGTGAACAGTTAAACCAAGCATTTGAGAAATCAATGACCAAG 660
QY 5201 ATCAAAAAACATTAATTAATTTGCCGAAATGCAATGAGAGAAAAGCGCTGAGTTG 5260
DB 661 CTGAATTCAGAGAGACCAATTAATGGAATAGTATCGACAGAGAGAGAAAAGAG 720
QY 5261 GAGCAACAGTCTCTATACAAATTTGAAAACAATCAGTTTATGCTATGTTAAATAACA 5320
DB 721 AAGCTTTAAACAAGTGTATACATTAATTAATCAATTCATTTTACATTAATTAATGTTA 780
QY 5321 GTAAATTTCAACGCGCAATGATCAAGATGAAAAATATCAATGATGATGCAAAAGATT 5380
DB 781 ATTAATTAATCAAGAGTTTATGATTAATTAAGACAAAACGATGAGCTATTAATAATCA 840
QY 5381 ATACTATGACCAATTAATGACAGTTCGAGTTGAGAGCAAAAAGAGCTCTGTGCAAG 5440
DB 841 AACCAATATCAACTATCAACCAACAGCAATTT-----AAATGAAATCACTATTCAC 891
QY 5441 GAGCTTCGCAAGTACTACTCTGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATA 5500
DB 892 TAGACCTCAACGCTGATTTAATTAAGAAATTAAGAGTCTACAGTTGAGAAAAGCCT 951
QY 5501 TTGACAAAGATTAGAGGAAGAAATTAATGAAATTAAGAAAAAGCAAAATGTTAATGTT 5560
DB 952 CGGCTATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1011
QY 5561 TAGCTAAAAATGCAAGTCAAGTGTCAAAATGCGACAGTCTTCGAGCAAGTGCAC 5620
DB 1012 AAACAATAGAGAGTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1071
QY 5621 AAGCTCAATGAGAGTGAAGTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 5680
DB 1072 TACCTGTTATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1131
QY 5681 TAAAAAATAGTACTCAAAATGTAAGAAATGCTTTGTTAAAGCAAAATCTCATTCATA 5740
DB 1132 TTCGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1191

QY 5741 TTAACAATTGGAATTGAGCTGAGTGTGAGAGAGCTGAGAGTGTG 5800
DB 1192 ATCAATTGACACAACTCTAACACAAAGAAATGTTCAATTGACCAAGATTAAAGAAATG 1251
QY 5801 TAGCGATGAATTAAGATTGTAATTAACGATAGCAAGATTAAATCAATGCAAAATGACTG 5860
DB 1252 AAGCTGTTAATGAGCTTAAGAAATAGCAAAATTCGAAATTTAATTAATTTCTGTCATG 1311
QY 5861 CGAAGGAAATGCGAGTTATTAACGAGTGTGAGCGGTAAATGCTAA 5909
DB 1312 CCAATTAAGGCTTAAGCTTAAGCAAGATTCAAGAACTTACAGATTAA 1360

RESULT 4

US-10-485-517-121
Sequence 121, Application US/10485517
Publication No. US2005025629A1
GENERAL INFORMATION:
APPLICANT: University of Sheffield
APPLICANT: Biosynex Incorporated
APPLICANT: Foster, Simon
APPLICANT: Mond, James
TITLE OF INVENTION: Antigenic Polypeptides
FILE REFERENCE: P100629MO
CURRENT APPLICATION NUMBER: US/10/485,517
CURRENT FILING DATE: 2004-02-02
PRIOR APPLICATION NUMBER: GB 0118825.9
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: GB 0200349.9
PRIOR FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 424
SOFTWARE: Patentin version 3.1
SEQ ID NO 121
LENGTH: 2189
ORGANISM: Staphylococcus aureus
US-10-485-517-121

Query Match 0.5%; Score 50.2; DB 1; Length 2189;
Best Local Similarity 46.4%; Pred. No. 1.6; Mismatches 188; Indels 0; Gaps 0;

Matches 163; Conservative 0; Mismatches 188; Indels 0; Gaps 0;
QY 2093 AACTTATGGAATGACGAGCTTTTATGAAACTATTAATATGCTTGTGACAA 2152
DB 343 AACCAATTATTAACGATGAGATTAATGAAAACGTAAGATTAATAAGCTGAAAAAT 402
QY 2153 TTGCGGATATGCTGATTTGTAATCGAAAGAAAAATGATATCAACAGTAAAAATAAA 2212
DB 403 TATTAGCGCATATTAAGCAAAACCTAGAAAGACAGTTTATGAAATTAATAAA 462
QY 2213 TGAATATTAATAATCCTTCAAAAATGCAAAAGCTGTTATGATTAATTAAGACTTTTA 2272
DB 463 TTCAAGATTAACCTTCTGAAAATTTAAAGCTGAGTACAGAAAGAAATTTAGAGATTACA 522
QY 2273 AGAGAGCTTTTGAAGAAAGAAAGCAAACTCCAGAAATATGATCGAAAGATTTGAATCTA 2332
DB 523 AGAAGCTTTAGATGAGCAAGTAATCAAGTATTAATTCGAATTCGAAATTTGACAAAGCA 582
QY 2333 TTGAAAAATTAATGATGCAATTTTCAAAAAATTTGATGAAAAACCGAGCTTTTACTTA 2392
DB 583 CAAATGAAAAATGACTGATTTTACAAAGATACAAAATATGTTTATGAAAGTGTGAGA 642
QY 2393 ATGAGAAAGATGACATTAATCTCCGAGATGAACTTCAAAAACAGAA 2443
DB 643 ATACGAATCTATGATGATGATCTTTGTTAAACACCTTATTAACAGGTA 693

RESULT 5

US-10-485-517-57
Sequence 57, Application US/10485517
Publication No. US2005025629A1
GENERAL INFORMATION:

APPLICANT: University of Sheffield
APPLICANT: Biosynex Incorporated
APPLICANT: Foster, Simon
APPLICANT: Mond, James
TITLE OF INVENTION: Antigenic Polypeptides
FILE REFERENCE: P100629MO
CURRENT APPLICATION NUMBER: US/10/485,517
CURRENT FILING DATE: 2004-02-02
PRIOR APPLICATION NUMBER: GB 0118825.9
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: GB 0200349.9
PRIOR FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 424
SOFTWARE: Patentin version 3.1
SEQ ID NO 57
LENGTH: 3174
TYPE: DNA
ORGANISM: Staphylococcus aureus
US-10-485-517-57

Query Match 0.5%; Score 50.2; DB 1; Length 3174;
Best Local Similarity 46.4%; Pred. No. 1.8; Mismatches 188; Indels 0; Gaps 0;

Matches 163; Conservative 0; Mismatches 188; Indels 0; Gaps 0;
QY 2093 AACTTATGGAATGACGAGCTTTTATGAAACTATTAATATGCTTGTGACAA 2152
DB 1298 AACCAATTATTAACGATGAGATTAATGAAAACGTAAGATTAATAAGCTGAAAAAT 1357
QY 2153 TTGCGGATATGCTGATTTGTAATCGAAAGAAAAATGATATCAACAGTAAAAATAAA 2212
DB 1358 TATTAGCGCATATTAAGCAAAACCTAGAAAGACAGTTTATGAAATTAATAAA 1417
QY 2213 TGAATATTAATAATCCTTCAAAAATGCAAAAGCTGTTATGATTAATTAAGACTTTTA 2272
DB 1418 TTCAAGATTAACCTTCTGAAAATTTAAAGCTGAGTACAGAAAGAAATTTAGAGATTACA 1477
QY 2273 AGAGAGCTTTTGAAGAAAGAAAGCAAACTCCAGAAATATGATCGAAAGATTTGAATCTA 2332
DB 1478 AGAAGCTTTAGATGAGCAAGTAATCAAGTATTAATTCGAAATTTGACAAAGCA 1537
QY 2333 TTGAAAAATTAATGATGCAATTTTCAAAAAATTTGATGAAAAACCGAGCTTTTACTTA 2392
DB 1538 CAAATGAAAAATGACTGATTTTACAAAGATACAAAATATGTTTATGAAAGTGTGAGA 1597
QY 2393 ATGAGAAAGATGACATTAATCTTCCGAGATGAACTTCAAAAACAGAA 2443
DB 1598 ATACGAATCTATGATGATGATCTTTGTTAAACACCTTATTAACAGGTA 1648

RESULT 6

US-10-485-517-12
Sequence 12, Application US/10485517
Publication No. US2005025629A1
GENERAL INFORMATION:
APPLICANT: University of Sheffield
APPLICANT: Biosynex Incorporated
APPLICANT: Foster, Simon
APPLICANT: Mond, James
TITLE OF INVENTION: Antigenic Polypeptides
FILE REFERENCE: P100629MO
CURRENT APPLICATION NUMBER: US/10/485,517
CURRENT FILING DATE: 2004-02-02
PRIOR APPLICATION NUMBER: GB 0118825.9
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: GB 0200349.9
PRIOR FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 424
SOFTWARE: Patentin version 3.1
SEQ ID NO 12
LENGTH: 3327
TYPE: DNA
ORGANISM: Staphylococcus aureus
US-10-485-517-12

Db 2339 CTCAGGCAATAGTCAACAAAGTCGACATAGACGAGATCAAAATGTCTACACAGAG 2398
Qy 857 TTGAATTCGATCTGTATTTGAAGCAGATGAAATTAATAAATTGTGTGAAGCTACAA 916
Db 2399 AAAAACAGAGCCCTTGAGAGCTTTAATCAAGAAACAAATGAGTCAATGATGAAATTC 2458
Qy 917 ATGGAGATTTTAAAGAAAGAGGAGAAAGAAACCTTAATACATCTCTTAAGTTTAT 976
Db 2459 AAGCAGCTTTAGCAAAATCAAAATGTTACAGACGAAAAAATATATATTAGAAACATTA 2518
Qy 977 CAGATGTGAAGCTTCCTGTAAGATTAATTAAGAAAGTCTTAGAAAGAAATGTTGACA 1036
Db 2519 GAAATGTGAACCTATGTAATTTGTA---AAACCAAGGCTATGAAATTAATAGAAAAA 2575
Qy 1037 TTACAGCTGAGCAAAAGATTTCTATGATGCAACTTTAGTCTAAGCTTGCAAA 1091
Db 2576 AAGCTGCGAACAACGACTTTAATTAATCAAAATCAAGATGACACCTAGAGAA 2630

RESULT 9
US-11-074-176-131

/ Sequence 131, Application US/11074176
/ Publication No. US20050250135A1
/ GENERAL INFORMATION:
/ APPLICANT: Kiehnhammer, Todd R.
/ APPLICANT: Russell, William M.
/ APPLICANT: Altermann, Eric
/ APPLICANT: McAniff, Olivia
/ APPLICANT: Perill, Andrea Azcarate
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding
/ TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
/ FILE REFERENCE: 5051-694
/ CURRENT APPLICATION NUMBER: US/11/074,176
/ PRIOR FILING DATE: 2005-03-07
/ PRIOR APPLICATION NUMBER: 60/551,161
/ PRIOR FILING DATE: 2004-03-08
/ NUMBER OF SEQ ID NOS: 381
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 131
/ LENGTH: 1293
/ TYPE: DNA
/ ORGANISM: Lactobacillus acidophilus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(1293)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: ORF 1295; Flay - Cell division protein
US-11-074-176-131

Query Match 0.5%; Score 45.6; DB 7; Length 1293;
Best Local Similarity 46.0%; Pred. No. 8.9; Indels 13; Gaps 4;
Matches 308; Conservative 0; Mismatches 349;

Qy 5502 TGACAAAGATTGAGAGAGAGAAATTAATGAAATTAAGAAAGCAATGTTAATGTTCT 5561
Db 39 TAAAAATGATGAT 98
Qy 5562 AGCTGAAATACGAGTCAAGTGTGTCACAAATGCGACAGTCTTCGAGCAAGTGACA 5621
Db 99 AAG 158
Qy 5622 AGCTGAGTAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5681
Db 159 TAATGATCTCAAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 218
Qy 5682 AAAAAATAGTACTCAAAATGTACGAAATGCTTTGTAAGAGAGAGAGAGAGAGAGAG 5741
Db 219 ATCGGTGAG 278
Qy 5742 TAAAAATGAGATTTGAG 5801

Db 279 AGAAACAGTTACTGAGAGAGTTGTGATGAGAGTCAATCTGAGAGAGAGAGAGAGAG 338
Qy 5802 AGCAGTGAATAGATTTGAAATTAATAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5861
Db 339 ATTAGAGCCGAG 395
Qy 5862 GAG 5921
Db 396 AAGAGCAATAG 452
Qy 5922 AGTGTCTGAG 5981
Db 453 AGTCATGAG 512
Qy 5982 ATCTACAAAG 6041
Db 513 TGAACAG 566
Qy 6042 TATTACTAG 6101
Db 567 TGCTAGTACAG 626
Qy 6102 TAAG 6161
Db 627 TAAGGTGAG 685
Qy 6162 TTGAGTACT 6171
Db 686 TTACTTAT 695

RESULT 10
US-10-485-517-11

/ Sequence 11, Application US/10485517
/ Publication No. US20050256299A1
/ GENERAL INFORMATION:
/ APPLICANT: University of Sheffield
/ APPLICANT: Biovexus Incorporated
/ APPLICANT: Foerster, Simon
/ APPLICANT: Mond, James
/ TITLE OF INVENTION: Antigenic Polypeptides
/ FILE REFERENCE: P1006290
/ CURRENT APPLICATION NUMBER: US/10/485,517
/ CURRENT FILING DATE: 2004-02-02
/ PRIOR APPLICATION NUMBER: GB 0118825.9
/ PRIOR FILING DATE: 2001-08-02
/ PRIOR APPLICATION NUMBER: GB 0200349.9
/ PRIOR FILING DATE: 2002-01-09
/ NUMBER OF SEQ ID NOS: 424
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 11
/ LENGTH: 2275
/ TYPE: DNA
/ ORGANISM: Staphylococcus aureus
US-10-485-517-11

Query Match 0.5%; Score 45.6; DB 1; Length 2275;
Best Local Similarity 43.4%; Pred. No. 11;
Matches 210; Conservative 0; Mismatches 274; Indels 0; Gaps 0;

Qy 7524 ATCAAG 7583
Db 786 AGCAGTAAATGAG 845
Qy 7584 TTTGTGAG 7643
Db 846 AGCAGTAAATGAG 905
Qy 7644 AACAGATATCAAGTCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7703
Db 906 GAAATTAATGAG 965

US-10-793-626-3712

Query Match 0.5%; Score 45.4; DB 1; Length 3180;
Best Local Similarity 49.0%; Pred. No. 13;
Matches 121; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 7119 AAATGTAACGGCAAAAGAGAAAAAATATTAAGCAAAACAGCAGCAAAATGCAAGATCGG 7178
DB 30 AGATCAAAAAGTAGAGCTTGTAATTAATGTTAAGAAAAATTAGTGAAGCTACTACACAAG 89
QY 7179 AGGAGCAGCAATCGGAGCCAAATGCTTGTAATTAATTTTGAACAAGCTGTGAAGATAG 7238
DB 90 TGAATTTCAATAGCTGAGCTGATGATGATGCTTATGCAATGATTGGACCGGAAAGA 149
QY 7239 AAAAAATTCGTAAGAAAAAGAAAGAAAGTTTAAAACTTTAGACGAAGTTAACAAAGA 7298
DB 150 TAAAGAAATTGGAAGAAATTTATAGAAAGCGCTTTTGAATATTCAGATATATGCTATTAAAGA 209
QY 7299 ACAAGATAAAAAAGTAAATGATGCTACGAAAAAAATCTTACAATGAGCAGATATTCTAC 7358
DB 210 AGGAGAACAAACTTTTCTGATCAATAGCAGAAAGATTAAATTAGCAGTTCACTTGC 269
QY 7359 AGAAGAT 7365
DB 270 CAATCAT 276

RESULT 14

US-10-793-626-3756/c
; Sequence 3756, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUJ480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3756
; LENGTH: 3650
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3756

Query Match 0.5%; Score 45.4; DB 1; Length 3650;
Best Local Similarity 49.0%; Pred. No. 13;
Matches 121; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 7119 AAATGTAACGGCAAAAGAGAAAAAATATTAAGCAAAACAGCAGCAAAATGCAAGATCGG 7178
DB 1975 AGATCAAAAAGTAGAGCTTGTAATTAATGTTAAGAAAAATTAGTGAAGCTACTACACAAG 1916
QY 7179 AGGAGCAGCAATCGGAGCCAAATGCTTGTAATTAATTTTGAACAAGCTGTGAAGATAG 7238
DB 1915 TGAATTTCAATAGCTGAGCTGATGATGATGCTTATGCAATGATTGGACCGGAAAGA 1856
QY 7239 AAAAAATTCGTAAGAAAAAGAAAGAAAGTTTAAAACTTTAGACGAAGTTAACAAAGA 7298
DB 1855 TAAAGAAATTGGAAGAAATTTATAGAAAGCGCTTTTGAATATTCAGATATATGCTATTAAAGA 1796
QY 7299 ACAAGATAAAAAAGTAAATGATGCTACGAAAAAAATCTTACAATGAGCAGATATTCTAC 7358
DB 1795 AGGAGAACAAACTTTTCTGATCAATAGCAGAAAGATTAAATTAGCAGTTCACTTGC 1736
QY 7359 AGAAGAT 7365
DB 1735 CAATCAT 1729

RESULT 15

US-10-793-626-2229
; Sequence 2229, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUJ480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2229
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2229

Query Match 0.5%; Score 45.2; DB 1; Length 570;
Best Local Similarity 47.3%; Pred. No. 8;
Matches 174; Conservative 0; Mismatches 188; Indels 6; Gaps 1;

QY 5862 GAAGGAAATGTCGAGTTATTACAGAGCTGATGCGGTAATTGCTAATTATGCAAGAAC 5921
DB 72 GAAAGACCGATATTAGATGTTATGCTCTGGGTAAATAATGATGTTAATTAAGTTAA 131
QY 5922 AGTGTGAGATGAGCCCTGAGCAATAGAGCCCTCAACAGTGAATGAATTAACAGG 5981
DB 132 AGACATGTTGTGTAAGCAAGACATAGGTGTACTAGCTTTGACAAAGCAAAAGACAT 191
QY 5982 ATCTCAAAAGCATATGTAAGAAATCTTCAAGTG-----ATTGCTAAAGAAAGAAACAGA 6035
DB 192 AGGTCAAAAAGCACTGTATTAAGCTAAGAGATGTGTCTAGCACTGTATTACAAGGATTTGG 251
QY 6036 TGAATTAATTACTACTCAAGGCAAGTAGATTAAGTGTAGATTAAGTATTCAAAATCT 6095
DB 252 AGATGTTTTGATATGATGAGGTCACTCATGAAAATTGGTAAATAAAGCTTTGAGAAAGT 311
QY 6096 TAAATTTACGAAGCTTATCACAAAAAGAAAAATAAGTAATAAAGATTGTTTAC 6155
DB 312 TGGTTTAACTTCAAGCTTTATGAAAAATGACCAATTCATTTGATTAATGACAGCTAT 371
QY 6156 CAATGATTCAGCTCTACTACTTTTAAATCTTTAATTGCAATGCGCGCTGTTCAAGACA 6215
DB 372 GATTAAAGAACTTTAAAAATGATATTAAGACTCTTTAATGAAGTTTAAGACTGTGACAG 431
QY 6216 AGCCGGAG 6223
DB 432 CGGTGAG 439

RESULT 16

US-10-793-626-3777/c
; Sequence 3777, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUJ480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3777

LENGTH: 2858
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3777

Query Match 0.5%; Score 45.2; DB 1; Length 2858;
Best Local Similarity 47.3%; Pred. No. 13;
Matches 174; Conservative 0; Mismatches 188; Indels 6; Gaps 1;

QY 5862 GAAGGAATGTCGAGTTATTACAGAGCTGATGCGGTAATGCTAATTAGCGAGAC 5921
DB 2138 GAAACACGGTATATTAAGATGTTATAGCTCTGTGTAAATAATGATTAAATGAAGTAA 2079
QY 5922 AGTGTGAGTGGCCCGGCGAGCAATAGAGCTCAACCACTGGAATGAATTAACAG 5981
DB 2078 AGACATTGGTGTAAAGCAAGACATAGGTGTACTACGTTGACAAAGCAAAAGACAT 2019
QY 5982 ATCTCAAAAGCATATGTAAAGATTCTACAGTG-----ATTGCTAAAGAAAGAACAG 6035
DB 2018 AGGTCAAAAGCACTTGAATTAAGCTAAGATGTGTACACTGTTATCAAGGTATTGG 1959
QY 6036 TGATTATATTACTACTCAAGGCGAAGTAGATAAGTGTAGATAAGTATTCAAAATCT 6095
DB 1958 AGATGTTTTGTATTTGTAGGTCTATCTATGAATTTGGTAATTAAGTCTTTGAGAAAGT 1899
QY 6096 TAAATTAACGAAGACTTATCACAAAAGAAATAAGTAAATAAAAGATTGTTAC 6155
DB 1898 TGGTTTAACTTACAGCTTATATGAAAATGCACTTACATTTGATTATACAGCTAT 1839
QY 6156 CAATAGTTCAGTACTCATCTTAAATCTTATTTGGCAATGCGCTGTTACAGACA 6215
DB 1838 GATTAAAGAACTTAAATAAGTATTAAGACTTCTTTATGAAGTTTAGACTGTGACAG 1779
QY 6216 AGCCGGAG 6223
DB 1778 CGGTGAG 1771

RESULT 17
US-10-793-626-3968/c
Sequence 3968, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 3968
LENGTH: 3002
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3968

Query Match 0.5%; Score 45.2; DB 1; Length 3002;
Best Local Similarity 48.4%; Pred. No. 14;
Matches 156; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

QY 7686 TGCCCTGAGATTAATTTCTAATGGAATCAATATCAGTATTTAAATTTCTAAGCT 7745
DB 1176 TACAAAGATGATTAAGAAATTTGTTCAATGATGATTAATTAACCTTACTACGCTCT 1117
QY 7746 ATTAGAAAAAATATTGATTGTTGTAAGTAATGCAATTGAGAGCGGAACAA 7805

DB 1116 TTACAGCAAAATATTATATTGTTGTACTGATTAACGAAAGAAAAATTCAGAGAG 1057
QY 7806 AGATTAAACGTAGAGCGGTAGCTGCCGAGCCATTATCTCA---AAGCAAGAAATGA 7862
DB 1056 AGGATATGATCTGTAATGTGTGTCAGACCACTCATTTAGCAAAATTCAAAAACGTTGA 997
QY 7863 AATGAATTCAGAGTTGAAAATTGAGACAGATATTTTCAATGAAGAAAATAGTAATTAG 7922
DB 996 AGATTAATTTAAGCGAATTGATTGATGAAATTAATTTGAAGGTAAAGAAATGAAT 937
QY 7923 CCTTCTAAGAAATTGGAAGAAATCAATGTCMAAGTGGAAAAAGAAACAGAGTAC 7982
DB 936 TGATCATATGTGTAAAGAAATTTAAGTATGATATTTATGAATTAACGCTAATAAAGAAAC 877
QY 7983 TGCTGAATCTCAGAGAGCTTCT 8004
DB 876 AACAGAAATCATTAATTAATTAAT 855

RESULT 18
US-10-793-626-3432/c
Sequence 3432, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 3432
LENGTH: 3063
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3432

Query Match 0.5%; Score 45.2; DB 1; Length 3063;
Best Local Similarity 47.3%; Pred. No. 14;
Matches 174; Conservative 0; Mismatches 188; Indels 6; Gaps 1;

QY 5862 GAAGGAATGTCGAGTTATTACAGAGCTGATGCGGTAATGCTAATTATGACGAAC 5921
DB 1889 GAAACACGGTATATTAAGATGTTATTAAGCTGTGTAAATAATGATGTTAATTAAGTAA 1830
QY 5922 AGTGTGAGTGGCCCGGCGAGCAATAGAGCTTCAACCACTGGAATGAATTAACAG 5981
DB 1829 AGACATTGGTGTAAAGCAAGACATAGGTGTACTACGTTTGAACAAAGCAAAAGACAT 1770
QY 5982 ATCTCAAAAGCATATGTAAAGATTCTACAGTG-----ATTGCTAAAGAAAGAACAG 6035
DB 1769 AGGTCAAAAGCACTTGAATTAAGCTAAGATGTGTACAGCTTATCAAGGTATTGG 1710
QY 6036 TGATTATATTACTACTCAAGGCGAAGTAGATAAGTGTAGATTAAGTAAAGTTCAAAATCT 6095
DB 1709 AGATGTTTTGTATTTGTAGGTCTATCTATGAATTTGGTAATTAAGTCTTTGAGAAAGT 1650
QY 6096 TAAATTAACGAAGCTTATCACAAAAGAAATAAGTAAATTAATTAATTAATGTTGTTAC 6155
DB 1649 TGGTTTAACTTACAGCTTATTAAGAAATGCAACATTTGATTTAATGACAGCTAT 1590
QY 6156 CAATAGTTCAGTACTCATCTTAAATCTTATTTGGAATTTGCCGCTGTTACAGACA 6215
DB 1589 GATTAAAGAACTTAAATAAGTATTAAGACTTCTTTATGAAGTTTAGACTGTGACAG 1530
QY 6216 AGCCGGAG 6223

Db 1529 CGGTGGAG 1522

RESULT 19

US-10-793-626-3676/C

; Sequence 3676; Application US/10793626
; Publication No. US20050255478A1

; GENERAL INFORMATION:

; APPLICANT: KIMBERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PUI480US

; CURRENT APPLICATION NUMBER: US/10/793,626

; CURRENT FILING DATE: 2004-03-04

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3676

; LENGTH: 3295

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-10-793-626-3676

Query Match 0.5%; Score 45.2; DB 1; Length 3295;

Best Local Similarity 48.4%; Pred. No. 14;

Matches 156; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

Db 7666 TGCCTATGCGAATTAAATTTCTAATGGAAGATCAATATCAGTATTAAATTTCTAAGCT 7745

Db 1325 TACAAAGATGATTAATAAGAAATGTTTCAATGATGTAATTAACCTACTACCGCTCT 1266

Qy 7746 ATTAGAAAAATATTTAGTATTTGTAATAAATATTCGAAATTCGAGCGGAGCAAA 7805

Db 1265 TTCAGAGCAAAATTAATTAATTTGTTTCTGATTAAGGCAAGAAAAATTCGAGAGA 1206

Qy 7806 AGGATTTAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAA---AAGCAAGAATGA 7862

Db 1205 AGGATATGATCCTGTAATATGTGTCTAGACCACTCTTAGAGCAATTCAAAAACGGTTGA 1146

Qy 7863 AATGATTCAGAGGTTGAATTTGAGAAGATATTTTCAATGGAAGAAATAGAGTAAGT 7922

Db 1145 AGATATATTTAAACGATTTGATTTAGATGGAATTAATTTGAAGTTAAAGAGTAACAT 1086

Qy 7923 CCCTCTTAAGGAATTTGGAAGGAATCAATGCTCAAGTGGAAGAAAAAGAAAGAGTAC 7982

Db 1085 TGATCATGATGTGTAAGAAATTTAAGTATGATTTATGAAATTTACGCTAAAAAGAAAC 1026

Qy 7983 TGCTGAATCTCAAGAGCTTCT 8004

Db 1025 AACGATCATTAATTAATTTAT 1004

RESULT 20

US-10-793-626-4138

; Sequence 4138; Application US/10793626

; Publication No. US20050255478A1

; GENERAL INFORMATION:

; APPLICANT: KIMBERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PUI480US

; CURRENT APPLICATION NUMBER: US/10/793,626

; CURRENT FILING DATE: 2004-03-04

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4138

; LENGTH: 3379

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-10-793-626-4138

Query Match 0.5%; Score 45.2; DB 1; Length 3295;

Best Local Similarity 48.4%; Pred. No. 14;

Matches 156; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4138

Query Match 0.5%; Score 45.2; DB 1; Length 3379;

Best Local Similarity 47.3%; Pred. No. 14;

Matches 174; Conservative 0; Mismatches 188; Indels 6; Gaps 1;

Qy 5862 GAAGGAAATGCGAGTATTATACAGAGTCTGATGCGGTAATGCTAATTAATGACAGAC 5921

Db 1261 GAAAGACGGTATATTAATGATTTAATGCTCTGCTGTATAAAATGATTTAAAGTAAA 1320

Qy 5922 AGTCTGAGTGGCCCGTGCAGCAATGAGACCTCAACAGTGTGAATGAATTAACAG 5981

Db 1321 AGACATTGTGTGAAGCAAGACATGAGTGTACTACGTTTGACAAAGCAAGAAACAT 1380

Qy 5982 ATCTCAAAAGCATATGTAAGATTTACAGTG----ATTGCTAAAGAAAGAAACGA 6035

Db 1381 AGTACAAAGACACTGTGATAAGCTAAAGATGTCTAGCACTGTTATCAAGGTAATTGG 1440

Qy 6036 TGATTATTAATCTACTCAAGGCAAGTAGATTAAGTGTAGATAAGTATTCAAAATCT 6095

Db 1441 AGATGTTTTTATATATGATGATCCTATGAAATTTGTAATAAGTCTTTGAGAAAGT 1500

Qy 6096 TAATATTAACGAGACTTATCACAAAAAAGAAATAAGTATTAAGAAATTTGTTAC 6155

Db 1501 TGGTTTAACTAGACCTTATGAAAAAATGACACATTAACATTTGATTTAATGACAGTAT 1560

Qy 6156 CAATGATTCAGTACTACTACTTTAAATCTTTAATGGCAAAATCCGCTGCTTACAGCA 6215

Db 1561 GATTAGAAAGCTTTAAATGATTTAAAGACTCTTTAATGAAGTTTGAAGCTGACAG 1620

Qy 6216 AGCGGAG 6223

Db 1621 CGGTGGAG 1628

RESULT 21

US-10-793-626-4114/C

; Sequence 4114; Application US/10793626

; Publication No. US20050255478A1

; GENERAL INFORMATION:

; APPLICANT: KIMBERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PUI480US

; CURRENT APPLICATION NUMBER: US/10/793,626

; CURRENT FILING DATE: 2004-03-04

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4114

; LENGTH: 3518

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-10-793-626-4114

Query Match 0.5%; Score 45.2; DB 1; Length 3518;

Best Local Similarity 48.4%; Pred. No. 14;

Matches 156; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

Qy 7666 TGCCTATGCGAATTAAATTTCTAATGGAAGATCAATATCAGTATTAAATTTCTAAGCT 7745

Db 3139 TACAAAGATGATTAATAAGAAATTTGATGATGTAATAATACTTACTACCGCTCT 3080

Qy 7746 ATTAGAAAAATATTTGATTTTATTTGTAATAAATTCGGAATTCGAGCGGAGCAAA 7805

Db 3079 TTCAGAGCAAAATTAATTAATTTGTTTACTGATTAAGCAAGAAAGAAAAATTCGAGAGA 3020

US-10-793-626-2905

Query Match	0.5%;	Score 45.2;	DB 1;	Length 4719;
Best Local Similarity	47.3%;	Pred. No. 16;		
Matches 174;	Conservative	0;	Mismatches 188;	Indels 6;
				Gaps 1

Qy	5862	GAAGGAAATATCCGAGTTATTTACAGAGCTGATGCGGTAATGCTAATTTATGCAAGAAC	5921
Db	3195	GAAAGACGGTATATTAGATGTTATTATGACTCTGGGTATAAAAATGATGTTAATAAGTAA	325
Qy	5922	AGTGTCTGAGTGGCCCGCTGCAGCAATAGAGCCTCAACCAAGTGTGAATGAAATTACAG	5981
Db	3255	AGACATTTGGTGGTAAAGCAAGAGCATAGTGGTACTACGTTTGAACAAGCAAAAGACAT	3314
Qy	5982	ATCTACAAAGCATATGTAAAAAGATTCTACAGTG-----ATTGCTAAAGAAAGAAACGA	6035
Db	3315	AGGTACAAAGACACTGTGATTAAGCTAAAGATGTGCTAGCACTGTATCAAGGGGTATGG	3374
Qy	6036	TGATTATATTACTACTCAAGGGCAAGTAGATTAAGTGGTAGTAAGTATTCAAAAATCT	6095
Db	3375	AGATGTTTTTATTTATTTATGTAAGTCATCCTATGAAATTTGTAAATTAAGCTTTGAAGAAAGT	3434
Qy	6096	TAATATTACAGAGCTTATCACAAGAAAAAGAAAAAATAGTAATAAAAAGAGATTGGTAC	6155
Db	3435	TGGTTTATCCTAGACTTATATGAAAAAATGCACCATTTACATTGATTTAATGACAGCTAT	3495
Qy	6156	CAATAGTTCAGCTACTCATTACTTTAAAACTTTAATGGCAATGCCGTGGTTCAGACA	6215
Db	3495	GATTAAAGAACTTTAAAAATGGTATTTAAAGACTTCTTATATGAAGTTTAGACTGACAG	3554
Qy	6216	AGCCGGAG 6223	
Db	3555	CGGTGGAG 3562	

RESULT 25
IIS-10-485

```

US-10-485-517-13
; Sequence 13, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; Applicant: University of Sheffield
; Applicant: Biosynexus Incorporated
; Applicant: Foster, Simon
; Applicant: Mord, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: PI066290
; CURRENT APPLICATION NUMBER: US/10/485,517
; PENDING FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-13

```

Query Match	0.5%	Score 45;	DB 1;	Length 2119;
Best Local Similarity	44.1%;	Pred. NO. 13;		
Matches 189; Conservative	0;	Mismatches 240;	Indels 0;	Gaps 0

QY	648	AAAGCAGAGAAATTACAGATTTTAAAAATTTTGTCAATATTTGTATCGAATTAATTTCTGG	707
Db	1686	AACTTCAGCTGTACTTCTTAATTAATAATGCAATTAAATGGTTATGAAAAATTGACAAATC	1744
QY	708	TCTGACCGGAGATTTTAAAGCTACCAAGACAAAATCTGGAGATATTATTTCTTTCAGCTCA	767
Db	1746	TAAACACATGCTGCACAGCGCATTTGATGCAATTTACACATTTTAAATTAATGACACAAAAAGC	1805
QY	768	CATGATTTCTCTCAAAAAGCTATGAGAAAAAATTCACATGTTGGAAAAAGAGATAGAAGA	827

Db
1806 AGATGTTAATCTAAATTAATGCTGCATCAATATTGCTGGCGTAATACTGTTAACA 1865

828 ATATGTAAGGAATACCAAGCAATATTGAATCTGATGCTGTATTGGAAGCAGATCG 887

Db
1866 ACAAGGTACAGATTTAAATACAGCGATGGGTA
CTTGCAAGGTGCAATCAATGATGACA 1925

888 AATATATAAATTAGTGGGAAGCTACCAATGGAGATTATTAAGCAAGGGAAGGGA 947

100
90
80
70
60
50
40
30
20
10
0

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

348 AGAAAC IATAACAC I C C T I AAG I T A CAGA I GAGAG C I CCG I AAGAG I AAA I AA 200

Db 1986 AAATGCCGTACAAGCTGCGAAGATATTTAAATAAATCAATGGTCAAAATAAAACGAA 2045

1008 AGGAAAGTCATAGGAAGAATGTTGACATTACAGCTGAAGCAAGAATTTCTATGATGC 1067

Db 2046 AGATCAAGTTACTGAAGCGATGAATCAAGTAATCTGCTAAATACTAGATGGTAC 2105

0Y 1068 AACCTTACT 1076

RESULT 26

```

US-10-793-626-4448/c
Sequence 4448, Application US/107933626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/10/793.626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4448
LENGTH: 2621
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4448

```

Query Match	0.54;	Score 45;	DB 1;	Length 2621;
Best Local Similarity	46.64;	Pred. NO. 14;		
Matches 144;	Conservative 0;	Mismatches 165;	Indels 0;	Gaps 0;

Qy	4373	ATGTCATTGCCAAAGATGTCAAAGAAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTAC	4432
Db	1956	ATTACATTGATTAAAGGATATCATATBAAGAGAGGTTATTCTACGCGGTCCAGGTAAC	1897
Qy	4433	TAAATGAAAAAGATTAATAAATATTATTAGAAGATCGTGTATTAAATACGACTGAAATGGTT	4492
Db	1896	CAGAGAGATTAGATTAAGAAAGATTTTTCGGGTCCACATATTTATTATATGATATATC	1833
Qy	4493	ATTATACGAAGGAACAATAGAAAAAGCAAAAAAGAAAGAAAGAGCGGTCAATTGTAAATG	4552
Db	1836	ACATGACGATTGCAACAAGAAAGAAATTTTCGACCTGTCAATCTGTAAATCACTTATATA	1777
Qy	4553	CTGCTTTATCGGTTCGCGAAGCGATTAATCCGTGAGAGAGTACATATTCAGTCATA	4612
Db	1776	ATCTTGACGAAGCATTTAGAAATGTCTATGATATCTAAATATATGTGCTTGCCTGATGTTA	1717
Qy	4613	CTGTTAAATAATTAATTTAAAGCAAGATTGATGAAACCAATTAAGAAAGCCGGAAGGATA	4672
Db	1716	TTGGTAAAGATTAAGATATCATTTGAGACATGTGCTCGTTCAATGGAAGCTGGTACAAATTG	1655
Qy	4673	AAATTCATG	4681

Db 1656 AAATTAATG 1648

RESULT 27

US-10-793-626-3353
; Sequence 3353, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3353
; LENGTH: 3589
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3353

Query Match 0.5%; Score 45; DB 1; Length 3589;

Best Local Similarity 46.6%; Pred. No. 16; Mismatches 165; Indels 0; Gaps 0;

Matches 144; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 4373 ATGTCATTGCCAAGATGTCAGAAAGAGTTCTGATCTAGCAAAAGAAATTCAGGCTTTAC 4432

Db 79 ATTACATTGATTAAGATATCAATGAAGAGAGAGTTATCTACGGTGTCCAGGTAAC 138

QY 4433 TAAATGAAAAGATTAATAATTTTGAAGATCGTGCTTATATGAGTCTGGAATGTT 4492

Db 139 CAGAAGAGATTAGATTAAGATATTTTGCCTGCTCACTATATTATTAATGTAATATC 198

QY 4493 ATTATACGAAGAACACTAGAAAAGAAAAGAAAAGAAAAGAGCGGCTATTGTAATG 4552

Db 199 ACATGACGATTCACACAGAAAGAAATTTTGGACCTCTCATGCTGTAATCACTTATATA 258

QY 4553 CTGCTTTATCGTTCGTGAGCGATTAATTCGCTGAGAGTAGTATTCAGTCAATA 4612

Db 259 ATCTTGACGAAGCTATTTGAATTTGCTAATGATATTAATATGCTGCTGGCTATGTTA 318

QY 4613 CTGTTAAATAATTAATTAAGCAATTTGAGTGAAGCAATTAAGAACCGGAGAGATA 4672

Db 319 TTGGTAAAGATTAAGATACATTGAGACATGTTGCTGTTCAATCGAAGCTGTACAAATTG 378

QY 4673 AAATTCATG 4681

Db 379 AAATTAATG 387

RESULT 28

US-10-793-626-835
; Sequence 835, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 835
; LENGTH: 1206
; TYPE: DNA
; ORGANISM: Artificial Sequence

FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-10-793-626-835

Query Match 0.5%; Score 44.8; DB 1; Length 1206;

Best Local Similarity 46.2%; Pred. No. 12; Mismatches 112; Indels 0; Gaps 0;

Matches 148; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 494 ATGTAAAGTTTAAATGAAATCATTCAGTAGATGAAAAGTAAATTCATTGAATC 553

Db 185 AAGTAATGATCATATCGGTGACAGTTCAGTTGCAAAAAACAATGATTAACCAAG 244

QY 554 CGAATGAAAGATTAACGTGTAAGAAAAATCAATGCTTGAAGCGATCGTTATAG 613

Db 245 CAATTCGAATTAATGACTGCTGCTCAAAATTCCTTGAGGCTGATGCCGTAATGTTG 304

QY 614 CGGCGATATTAGATTGAAGATCTGCAATCTAAGACAGAAATTCAGATTTTAA 673

Db 305 AACAACTATTGAATCTGAAACACTTTTACAAATTAAGAAAATCCTTAAACATTAGAA 364

QY 674 ATTTAGTCAATTAATGATGATGAAATTAATTTGCTGACCGAGATTTAAAGCTACCA 733

Db 365 ATATTGCTACAAAGTGAAGAAATTAAGCTGATATTTGATACATATAAGGATAGC 424

QY 734 AGACAAATCTGAGATTTATTTCTTCACTCACATGATTTCTCAAAAAGCTATG 793

Db 425 GTATTACTCAGTGTGATGACGCTTACGATACATACGTTATCTAAAGTCGAGTGG 484

QY 794 GAAAAATTTCACTGTGGA 813

Db 485 CTCGAAAACCAACTGTGCA 504

RESULT 29

US-10-793-626-4188
; Sequence 4188, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4188
; LENGTH: 3354
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4188

Query Match 0.5%; Score 44.8; DB 1; Length 3354;

Best Local Similarity 46.2%; Pred. No. 17; Mismatches 172; Indels 0; Gaps 0;

Matches 148; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 494 ATGTAAAGTTTAAATGAAATCATTCAGTAGATGAAAAGTAAATTCATTGAATC 553

Db 964 AAGTAATGATCATATCGGTGACAGTTCAGTTGCAAAAAACAATGATTAACCAAG 1023

QY 554 CGAATGAAAGATTAACGTGTAAGAAAAATCAATGCTTGAAGCGATCGTTATAG 613

Db 1024 CAATTCGAATTAATGACTGCTGCTCAAAATTCCTTGAGGCTGATGCCGTAATGTTG 1083

QY 614 CGGCGATATTAGATTGAAGATCTGCAATCTAAGACAGAAATTCAGATTTTAA 673

Db 1084 AACAACTATTGAATCTGAAACACTTTTACAAATTAAGAAAATCCTTAAACATTAGAA 1143

QY 674 ATTGATCAATATAGTATGCAATTAATTCGTCTGACCGAGATTAAAGCTACCA 733
DB 1144 ATATTCGCTACAGGTGAAGAAATTAAGCTGGATATTTGATACATAAAGGTATGC 1203
QY 734 AGACAAATCTGAGATATATTTCTTTCAGCTCAGCTATATCTCTCAAAAAGCTATGG 793
DB 1204 GTATTACTCAGGTGATAGACAGCTTGTATACATACGCTTATTAAGTGCAGATGG 1263
QY 794 GAAAAAATCACTGTTGGA 813
DB 1264 CTCGAAAAACCAACTGTTGCA 1283

RESULT 30

US-10-793-626-4025/c
; Sequence 4025, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4025
; LENGTH: 3497
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4025

Query Match 0.5%; Score 44.8; DB 1; Length 3497;
Best Local Similarity 46.2%; Pred. No. 17;
Matches 148; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 494 ATGCTAAATTTTATGCAATCATTCGCTAGATGAGAAAGTAAATTCATTGAATC 553
DB 3236 AAGTAATATCATATACGGGTGACGGTTCAGTTCAGAAAAACAATTGATTAACCAAG 3237
QY 554 CGAATGGAAGCATTCACGTAGAGAAAGAAATCAATCTGTTGAAGCATCGTTATATG 613
DB 3236 CAATTGAAATATGACTGCTGCTCAAAATCTCTGAGCTATGCGGTAGTAATGTTG 3177
QY 614 CGCGGATATTAAGATTGAAAGATCTGCAATACTAAAGACAGAAATTACAGATTTTAAA 673
DB 3176 AACAAACATTAATCTGAAACAACTTTACAAATTGAAAAATCCTTAAACATTAGAAA 3117
QY 674 ATTTAGTCAATATTAAGTATGCAATTAATTCGTGCTGACCGAGATTAAAGCTACCA 733
DB 3116 ATATTCGCTACAGGTGAAGAAATTAAGCTGGATATTTGATACATAAAGGTATGC 3057
QY 734 AGACAAATCTGAGATATATTTCTTTCAGCTCAGCTACATAGATTCCTCAAAAAGCTATGG 793
DB 3056 GTATTACTCAGGTGATAGACAGCTTGTATACATACGCTTATTAAGTGCAGATGG 2997
QY 794 GAAAAAATCACTGTTGGA 813
DB 2996 CTCGAAAAACCAACTGTTGCA 2977

RESULT 31

US-10-793-626-3818
; Sequence 3818, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US

; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3818
; LENGTH: 3066
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3818

Query Match 0.5%; Score 44.6; DB 1; Length 3066;
Best Local Similarity 51.8%; Pred. No. 17;
Matches 101; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 7183 GCAGCAATCGAGCCAAATGCTTGTAAATATTTTGAACAGCTGTAGAAATGAAAA 7242
DB 2153 GTAGAACTGGAAGACAAATCAGAAATGATCAAAATATTTGAAAATCTGATTTCTAACACAGAT 2212
QY 7243 AATTCTGAAGGAAAGGAAACGAACTTTTAAACTTTAGACGAAATTAAACAAGACAA 7302
DB 2213 AATAGTGAAGCTAAATGCTAGCGATTCAAGAAATAATTCAGAAAGATATCAAGATGAA 2272
QY 7303 GATAAAAAGTAAATGATGCTACGAAAAAAATCTTACAATGACAGGTATTTCTACAGA 7362
DB 2273 GAGTCAGAGCTCAAGATCTAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 2332
QY 7363 GATACCTCTGTAAAA 7377
DB 2333 GAAAAATATTTTAAAG 2347

RESULT 32

US-10-793-626-3849
; Sequence 3849, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3849
; LENGTH: 3170
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3849

Query Match 0.5%; Score 44.6; DB 1; Length 3170;
Best Local Similarity 44.6%; Pred. No. 18;
Matches 271; Conservative 0; Mismatches 324; Indels 12; Gaps 2;

QY 5983 TCTACAAAAGCATATGTAAGATTTCTACAGGATGCTTAAAGAGAAAGATGATAT 6042
DB 1185 TTTACACATGCAATGGAAGAAATTTAAAGAAATAGCTACAGTACATTTGATTTTCA 1244
QY 6043 ATTTACTCAAGGCAAGTAAAGTGTAGATTAAGTATTCATTAATCTTATATTT 6102
DB 1245 ATTTAAGATACCAACATCAAGATGATTTGCTTTAAATATTTGACATTTGATAT 1304
QY 6103 AACGAAGCTTATCACAAAAAAGAAAAATTAAGTATTAATAAAGATTTGTATCAATAGT 6162

```
Db      1305 CAACACGATGTCGCCCAAAAGTGAAGTCAATTGAACAAAATACTATTAAGT 1364
Qy      6163 TCAGCTACTCATCTTTAAATCTTTATTTGGCAATGCCGCTGTTCAAGACAGCCGA 6222
Db      1365 GGAA-----ATGTTTAAAGTCTTAGTTCAAAATGACATTGAGAGAAATCATTTTGAAT 1418
Qy      6223 GTGGAGGAACTGTTAATATCAACAAGTTTATGGAAGAAAGAAAGCTTTGAGAAAT 6282
Db      1419 AAGGCACTTAATGTAAGTTT-----GGTTAAAGCTTTTAAAGCTGCTTGAATTT 1472
Qy      6283 TCTATATTAATGCAAAACATTATCTGTAATAATCAGAGATTACAGAAATCAATCGA 6342
Db      1473 GATTAAGTCTTTTGAACAGATTCAAGAAATCAAGATGATGACCTTGCAATCGTTAGA 1532
Qy      6343 GTAGTAGGTTCTGTTGGTGTGGTGAATGTAGAGATGAGACTTCTTCTGATACCAAT 6402
Db      1533 GCACATATTCAACAAGAAATGATCAAAAGTGCAGAGAACTGAAATAATTGAAAAA 1592
Qy      6403 ATTATTAAGAAATATCCAAAGACAGAGTTGGAATAATCAATGTCGATGAAGTTTC 6462
Db      1593 ATGAAAGCAGAAAGCCGAAACAAAGATTAATTAAGAAAGTACAGTGAATAATGTCAG 1652
Qy      6463 GGAGAAAGAGCTGAATTTACAGCAGATTCTTAAGCAAGAAATTTCTTTTGGAGTCGA 6522
Db      1653 ATTTGAAACCAATTCAAGTTGAAATATTAACCAATTGATCAATTTGAAGAAAGAA 1712
Qy      6523 GTTCGACAGCCGGGGTAGAGCCGAGTGCAGAGAACCGTTTCCGTAATCAATTGCA 6582
Db      1713 TTCAAGTAGCTATTGAAGTGTATATTTGATATTAACCTTAAGAACTTAAGTAGGA 1772
Qy      6583 GGAAAGA 6589
Db      1773 CGTCATA 1779
```

```
RESULT 33
US-10-793-626-2051
; Sequence 2051, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2051
; LENGTH: 4329
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2051
```

```
Query Match      0.5%; Score 44.6; DB 1; Length 4329;
Best Local Similarity 44.6%; Pred. No. 19;
Matches 271; Conservative 0; Mismatches 324; Indels 12; Gaps 2;

Qy      5983 TCTACAAAGCATATGTAAGATTCTACAGTATGCTTAAGAAAGAAAGATGATAT 6042
Db      196 TTTACACATGCAATTGAAGAAAGATTAAAGAAATAGCTACAGTACGAAATTTTCA 255
Qy      6043 ATTAATACTCAAGGCGCAAGTATGAAGTGTAGATTAAGTATTTCAAAATCTTAATAT 6102
Db      256 ATTAAGATACCAACATCAAGTGAAGTGTGCTTTAAATAATATTTGCGACATTTGATGAT 315
Qy      6103 AACGAAGCTTATCACAAAAGAAATAATAGTATATAAAAGATTTGTTACCAATAGT 6162
Db      316 CAACACGATTTGTCGCAAAAGTGAAGGTCAATTTGAACAAAAAAACATCATTAAGT 375
```

```
Qy      6163 TCAGCTACTCATCTTTAAATCTTTATTTGGCAATGCCGCTGTTCAAGACAGCCGA 6222
Db      376 GGAA-----ATGTTTAAAGTCTTAGTTCAAAATGACATTGAGAGAAATCATTTTGAAT 429
Qy      6223 GTGGAGGAACTGTTAATATCAACAAGTTTATGGAAGAAAGAAAGCTTTGAGAAAT 6282
Db      430 AAGGCACTTAATGTAAGTTT-----GGTTAAAGCTTTTAAAGCTGCTTGAATTT 483
Qy      6283 TCTATATTAATGCAAAACATTATCTGTAATAATCAGAGATTACAGAAATCAATCGA 6342
Db      484 GATTAAGTCTTTTGAACAGATTCAACAAATCAAGATGATGACCTTGCAATCGTTAGA 543
Qy      6343 GTAGTAGGTTCTGTTGGTGTGGTGAATGTAGAGATGAGACTTCTTCTGATACCAAT 6402
Db      544 GCACATATTCAACAAGAAATGATCAAAAGTGCAGAGAACTGAAATAATTGAAAAA 603
Qy      6403 ATTATTAAGAAATATCCAAAGACAGAGTTGGAATAATCAATGTCGATGAAGTTTC 6462
Db      604 ATGAAAGCAGAAAGCCGAAACAAAGATTAATTAAGAAAGTACAGTGAATAATGTCAG 663
Qy      6463 GGAGAAAGAGCTGAATTTACAGCAGATTCTTAAGCAAGAAATTTCTTTTGGAGTCGA 6522
Db      664 ATTTGAAACCAATTCAAGTTGAAATATTAACCAATTGATCAATTTGAAGAAAGAA 723
Qy      6523 GTTCGACAGCCGGGGTAGAGCCGAGTGCAGAGAACCGTTTCCGTAATCAATTGCA 6582
Db      724 TTCAAGTAGCTATTGAAGTGTATATTTGATATTAACCTTAAGAACTTAAGTAGGA 783
Qy      6583 GGAAAGA 6589
Db      784 CGTCATA 790
```

```
RESULT 34
US-10-793-626-1557/c
; Sequence 1557, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1557
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1557
```

```
Query Match      0.5%; Score 44.4; DB 1; Length 1500;
Best Local Similarity 52.7%; Pred. No. 15;
Matches 96; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy      7134 AGAGGAAAAAATTTTAAGCAACAGCAATTCAGAGAAATGCGAGAGACCAATCGG 7193
Db      699 AGAGAAATTAAGCTTAAGCTTAACATGACAGAGTGTGAATGATATTCAGCTAATTTCTGG 640
Qy      7194 AGCCAATGCTTGGTAAATTAATTTTGAACAGCTGTAGAGATAGAAAAAATTTCTGAGG 7253
Db      639 ACCATTAAACACTGTAAGTAAATCTTGTAAACTGTATAGTACAGAAACAAACAAGAT 580
Qy      7254 AAAAGAACAGAACTTTTAAAACTTTTGAAGCAAGTTTAACAAAGAACACATATAAAAGT 7313
Db      579 TGAAGGTAAGTTTCTTTAATTTCTTTGAACCATCATCTACGATTAAGATTAAGTAAGG 520
```

QY 7314 AA 7315
DB 519 AA 518

RESULT 35

US-10-793-626-4198/c
; Sequence 4198, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4198
; LENGTH: 3145
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4198

Query Match 0.5%; Score 44.4; DB 1; Length 3145;

Best Local Similarity 45.4%; Pred. No. 19; Mismatches 191; Indels 0; Gaps 0;
Matches 159; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 4967 CTGCTTGTATTAAGGAAAGGTAATCTTTAGTGAAGAATACAAAGTCAATG 5026
DB 367 CTAGATATTAAGAAAGTAACTCGGAGAGATATTCAGAAATTAATGAATACCTA 308
QY 5027 TACAAGCTTGAATGATTTCTCATATTAACAAGTTTCTGCTGAGCGCTGCAAGTATTA 5086
DB 307 AAGATTAATAAATGAAGATTTCTTCTTAATATGCTATGTTTAAACAAAAATGCTTA 248
QY 5087 AGCAGGCTGAATCGAGGAATGATCTGTCAATCGTGTCTGATGAAGAAAGGACTT 5146
DB 247 AGCGTATTCAGCTGAATGATTCGCAAGCTATTTGATCAATTAAGTACAAAGCAAGC 188
QY 5147 TAGTTAGTATTTCTGAGTTTGAAGAGTAAGTTCTTTCAATGATGAAGAAAGTCAAA 5206
DB 187 AGCTTAACAGGTTGATTTGCTGGGCCAGTTTCATCACTTTATCTGATATCAAT 128
QY 5207 AAACATAATATCAATTCGCGAAATGCAATGAGAGAAAGCGGCTGAGTTGAGCA 5266
DB 127 ATTAAACAGACATTTATTCAGAGAGCTATCAGAAAGGTGATGATTTGATGCAAC 68
QY 5267 CAGTTGCTATCAATATTTGAAAAATCAATCAGTTATGCTATTTGTAATAA 5316
DB 67 AATTAAGATTAAGTAACTTTATTAAGATTAATGATATGATCAATCAACA 18

RESULT 36

US-10-793-626-4195/c
; Sequence 4195, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4195
; LENGTH: 3328

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4195

Query Match 0.5%; Score 44.4; DB 1; Length 3328;
Best Local Similarity 52.7%; Pred. No. 19; Mismatches 96; Indels 0; Gaps 0;
Matches 96; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 7134 AGAGAAAAAATATTAAGCAACGACGCAATGACGAATCGAGAGCGCAATGG 7193
DB 2236 AGAGAAATTAAGCTATGCTAACTAGAGCAAGTATGAAATGATTCGATTAATTCGG 2237
QY 7134 AGCAATGCTTGGTAATTAATTTGAAACAGCTGTAGAAGTAAAGAAATTCGAAG 7253
DB 2236 ACCATTAAACACTGTAAGTAACTCTTGTAAACGTAATGACAGAAACAAAGAAAT 2177
QY 7254 AAAAGAACAGAGTTTAAATTAATTTAGAGAGTTTAAACAAAGAAAGATTAAGT 7313
DB 2176 TGAAGTAAAGTTTCTTTAATTTCTTTGAAACCATCTGATTAAGATTAAGTAAAG 2117
QY 7314 AA 7315
DB 2116 AA 2115

RESULT 37

US-10-793-626-3503/c
; Sequence 3503, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3503
; LENGTH: 3335
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3503

Query Match 0.5%; Score 44.4; DB 1; Length 3335;
Best Local Similarity 49.6%; Pred. No. 19; Mismatches 114; Indels 0; Gaps 0;
Matches 114; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 5943 AGCAATAGAGCCCTCAACCAAGTGAATGAATTAAGATCTACAAAGATATGTAA 6002
DB 2327 AGCTAACCAACTGTGATGATTTTACATTCACGTAACAAAAATATGATGCACTTAA 2268
QY 6003 AGATTCTACAGTATGCTTAAGAAAGAAACAGATATTAATTTACTACTCAAGGCAAGT 6062
DB 2267 ACAAGCAATGTGATTTGATGCGCAGTAGACCAACTGTTAGTTACTTAAAGAAATGT 2208
QY 6063 AGATTAAGTGTAGTAAAGATTTCAAAAATCTTAATTTTAAAGAAAGCTTATCAAAA 6122
DB 2207 CAAAAAGCGCAGTAATTAATGATGTTGTAATACCTATGAAATGAAAAATTTAA 2148
QY 6123 AAGAAAAATTAAGTAAATTAAGATTTGTTACCAATAGTCAAGTACTC 6172
DB 2147 AGGAGTGTACATATGATGAAGTAAAGAAATAGCTAGTGCATTAATCTC 2098

RESULT 38

```
US-10-793-626-4066/c
; Sequence 4066, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4066
; LENGTH: 3771
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4066

Query Match          0.5%; Score 44.4; DB 1; Length 3771;
Best Local Similarity 45.4%; Pred. No. 20;
Matches 159; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 4967 CTGCTTGATTAACAGAAAGGAAATCCCTTTAGTGAAGAAATACAAAGTCAATG 5026
DB 2332 CTAGATATTGAAGAAAGTAACTCGCAGAGATATTCCAGAAATTAATAATGAAATACCTA 2273
QY 5027 TACAAAGCTTGAATGATTTCTATTTACAAAGCTTTCTGCTGAGAGCGCTGCAAGATTA 5086
DB 2272 AAGATACATAAAATGAGATTTCTTTCTAATATGCTATGTTTAAACAAAATTTGCTA 2213
QY 5087 AGCAGGCTGGAATCGAGGAATGATGTTCTGCAATGCTGCTGATGAACGGAAGCTT 5146
DB 2212 AGCGTAATCAGTGAATTTGCGCAAGCTATTGTGATCATTTAGATACAGCAAGCAGC 2153
QY 5147 TAGTAGTATTTCTGAGTTTGAAGAGTAAAGTTCTTTCAATGTAGATGCAAAAGATCAA 5206
DB 2152 AGGTTAAACAGGTGATTTGCTGGGCCAGGTTTCACTTAATCTGATTAATCAAT 2093
QY 5207 AAACATAATTAATCAATTTGCCGGAATGCAATGAGAGAAAGCGGTGAGTTGAGCA 5266
DB 2092 ATTTAACAGACATTTATTCAGAAAGCTATCAAAAGGTGATGATTTGATATGCAACAC 2033
QY 5267 CAGTGTCTATCAATATTTGGAAGAAACATCACTTATAGCTATTGTAATA 5316
DB 2032 AATCTAAGAAATCTAATTTTATTTAGATATGATATCAGCTATCCACA 1983

RESULT 39
US-10-793-626-49
; Sequence 49, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 2454
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-49
```

```
Query Match          0.5%; Score 44.2; DB 1; Length 2454;
Best Local Similarity 48.6%; Pred. No. 19;
Matches 152; Conservative 0; Mismatches 158; Indels 3; Gaps 1;

QY 7686 TGCCATATCAGAAATTAATTTCTAATGAGATCAATATCATGATTAATAAATTTCTAAGT 7745
DB 2142 TACAAAAGATGATTAATAAAGAAATTTGTTACATGATGTAATTAACCTTACTACCGCT 2201
QY 7746 ATTAGAAAAAATTTGATGTTATTTGTAAGATTAATTCGAATTCGAGACCGAAGCAAA 7805
DB 2202 TTCAGAGCAAAATATTAATATTTGTTGTTACTGATTAACGAAAGAAAAAATTCGAGAA 2261
QY 7806 AGGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAA--AAGCAAGATGA 7862
DB 2262 AGGATATGATCTCGTAATATGTTGCTAGACCACTCATTAAGCAATTCAAAAACGTTGA 2321
QY 7863 AATGAATTCAGAGGTGAATTTGAGAGATGATTTTCAATGATGAAGAAAAATGAGTAAGT 7922
DB 2322 AGATTAATTTAAGCGAATTTGATTTAGATGAATTAATTTGAAGTTAAAGAGTAACAT 2381
QY 7923 CCTTCTAAGGAATTGGAAGAGAAATCAATGCTCAAGTGAAGAAAAAGAAAGAGTAC 7982
DB 2382 TGATCATGATGTTAAGAAATTTAAGTATGATATTTATGAATAATTACAGCTAAAAAGAAC 2441
QY 7983 TGCTGAATCTCA 7995
DB 2442 AAGCAATCATTA 2454
```

```
RESULT 40
US-10-793-626-1527
; Sequence 1527, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1527
; LENGTH: 2454
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-1527

Query Match          0.5%; Score 44.2; DB 1; Length 2454;
Best Local Similarity 48.6%; Pred. No. 19;
Matches 152; Conservative 0; Mismatches 158; Indels 3; Gaps 1;

QY 7686 TGCCATATCAGAAATTAATTTCTAATGAGATCAATATCATGATTAATAAATTTCTAAGT 7745
DB 2142 TACAAAAGATGATTAATAAAGAAATTTGTTACATGATGTAATTAACCTTACTACCGCT 2201
QY 7746 ATTAGAAAAAATTTGATGTTATTTGTAAGATTAATTCGAATTCGAGACCGAAGCAAA 7805
DB 2202 TTCAGAGCAAAATATTAATATTTGTTGTTACTGATTAACGAAAGAAAAAATTCGAGAA 2261
QY 7806 AGGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAA--AAGCAAGATGA 7862
DB 2262 AGGATATGATCTCGTAATATGTTGCTAGACCACTCATTAAGCAATTCAAAAACGTTGA 2321
QY 7863 AATGAATTCAGAGGTGAATTTGAGAGATGATTTTCAATGATGAAGAAAAATGAGTAAGT 7922
DB 2322 AGATTAATTTAAGCGAATTTGATTTAGATGAATTAATTTGAAGTTAAAGAGTAACAT 2381
QY 7923 CCTTCTAAGGAATTGGAAGAGAAATCAATGCTCAAGTGAAGAAAAAGAAAGAGTAC 7982
DB 2382 TGATCATGATGTTAAGAAATTTAAGTATGATATTTATGAATAATTACAGCTAAAAAGAAC 2441
QY 7983 TGCTGAATCTCA 7995
DB 2442 AAGCAATCATTA 2454
```

Qy	7923	CCCTTCTAAAGGAAATTGGAAGAAATCAATGTC	CAAAAGTGGAAAAAGAAACAGAGTGAC	7982
Db	2382	TGATCATGATGGTAAAGAAATTTAAGTA	TATATTATGAATTACAGCTAAAAAGAAC	2441
Qy	7983	TGCTGAATCTCAA	7995	
Db	2442	AACAGAAATCATTA	2454	

Search completed: November 26, 2005, 06:38:49
 Job time : 317 secs